

**UNIVERSIDADE FEDERAL DA PARAÍBA  
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**STÉLIO COELHO LOMBARDI FILHO**

**ENSAIOS SOBRE EDUCAÇÃO NA REDE PÚBLICA DE ENSINO FUNDAMENTAL:  
ANÁLISES A PARTIR DOS MICRODADOS DA PROVA BRASIL**

**JOÃO PESSOA – PB  
2015**

**STÉLIO COELHO LOMBARDI FILHO**

**ENSAIOS SOBRE EDUCAÇÃO NA REDE PÚBLICA DE ENSINO FUNDAMENTAL:  
ANÁLISES A PARTIR DOS MICRODADOS DA PROVA BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba – UFPB, em cumprimento às exigências do Curso de Mestrado em Economia.

**Orientadora:** Prof<sup>ª</sup>. Dra. Mércia Santos da Cruz

**JOÃO PESSOA – PB  
2015**

L842e      Lombardi Filho, Stélio Coêlho.  
              Ensaio sobre educação na rede pública de ensino  
              fundamental: análises a partir dos Microdados da Prova Brasil /  
              Stélio Coêlho Lombardi Filho.—João Pessoa, 2015.  
              76f.  
              Orientadora: Mércia Santos da Cruz  
              Dissertação (Mestrado) – UFPB/CCSA  
              1. Educação. 2. Educação – ensino fundamental – rede  
              pública. 3. Desigualdade de contingência. 4. Avaliação de  
              impacto. 5. Mais Educação.

UFPB/BC

CDU: 37(043)

**STÉLIO COÊLHO LOMBARDI FILHO**

**ENSAIOS SOBRE EDUCAÇÃO NA REDE PÚBLICA DE ENSINO FUNDAMENTAL:  
ANÁLISES A PARTIR DOS MICRODADOS DA PROVA BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba – UFPB, como requisito para a conclusão do curso de Mestrado em Economia.

Submetida à apreciação de banca examinadora, sendo aprovada em: \_\_\_\_/\_\_\_\_/\_\_\_\_

**BANCA EXAMINADORA**

---

**Prof.<sup>a</sup> Dra. Mércia Santos da Cruz**  
Departamento de Economia / PPGE – UFPB  
Orientador

---

**Prof. Dr. Paulo Aguiar do Monte**  
Departamento de Economia / PPGE – UFPB  
Examinador Interno

---

**Prof. Dr. Guilherme Diniz Irffi**  
Departamento de Economia Aplicada / CAEN – UFC  
Examinador Externo

*A Deus.*

*Aos meus pais, Stélio e Luíza.*

*Às minhas irmãs, Andréa e Juliana.*

*Aos meus amigos.*

## AGRADECIMENTOS

*Agradeço primeiramente a Deus, por sempre iluminar meu caminho e me dar forças para não desistir dos meus sonhos. Muito obrigado, Senhor, por tudo que tens feito por mim.*

*Aos meus pais, Stélio e Luíza, pelo amor incondicional e pelo apoio mais que necessário para a conclusão dessa etapa.*

*Às minhas irmãs, Andréa e Juliana, por sempre acreditarem em mim e me apoiarem em todas as minhas escolhas.*

*Agradeço imensamente à minha orientadora, a Prof.<sup>a</sup> Dra. Mércia Santos da Cruz, pelos conhecimentos transmitidos e pela dedicação, responsabilidade e paciência apresentadas durante toda a construção dessa dissertação. Serei eternamente grato.*

*Ao Prof. Dr. Guilherme Diniz Irffî, por toda a ajuda fornecida e pelo valioso aprendizado durante esses últimos dois anos.*

*Aos docentes do Departamento de Economia da UFPB, em especial aos professores Hilton Martins, Ignácio Tavares, Aléssio Tony, Paulo Aguiar, Adriano Paixão, Magno Batista, Liedje Siqueira, Edilean Kleber, Hélio Ramos, Maria da Conceição, Ivan Targino, Luiz Renato, Sinézio Maia, Alysson André, Paulo Fernando e Alexandre Martins. Muito obrigado por todos os ensinamentos e pela inestimável contribuição de cada um à minha vida acadêmica e profissional.*

*Ao coordenador do Programa de Pós-Graduação em Economia, o Prof. Dr. José Luís da Silva Netto Júnior, pela competência e responsabilidade.*

*Aos funcionários do PPGE, Risomar Farias e Ricardo Cataldi, pela amizade e boa vontade sempre demonstradas. Sem vocês a jornada teria sido muito mais difícil.*

*Ao amigo e professor da UFMT, Fábio Nobuo Nishimura, pelas valiosas sugestões.*

*Aos meus amigos, Murilo Callou, Leonardo Fagundes, Eduardo Cavalcanti, Jéssica Callou, Thaysa Cristina, Débora Luz, Veluma Wanderley, Pedro Borges, Lucas Oliveira, Thaís*

*Terehoff, Thomaz Choaíry, Narjara Saraiva, Clarissa de Barros, Victor Gomes, Fabrícia Carvalho, Camila Mirella, Thiago Fauvrelle e Vitor Morosine, pelos conselhos, apoio e momentos de descontração.*

*Aos amigos e colegas de turma, Allyne de Almeida, Otoniel Rodrigues, Celina Santos, Alexandre Loures, Sérgio Oliveira, Julyan Gleyvison, Patrícia Carvalho, Bruno Moura, Priscila Rodrigues e Juliane Ciríaco, pelos conselhos valiosos e pelos momentos de estudo e diversão.*

*Por fim, agradeço à CAPES pelo apoio financeiro fornecido durante todo o curso de Mestrado.*

*Se a educação sozinha não transformar a sociedade,  
sem ela tampouco a sociedade muda.*

*(Paulo Freire)*

*A educação é a arma mais poderosa que você pode  
usar para mudar o mundo.*

*(Nelson Mandela)*

*O Senhor é meu pastor e nada me faltará.*

*(Salmos: 23, 1-2)*



## RESUMO

Esta dissertação é composta por dois ensaios que buscam responder duas questões centrais. No primeiro ensaio o objetivo principal é verificar a ocorrência de desigualdade de contingência, no sentido descrito por Rawls (1971), entre os estudantes do 5º e 9º ano da rede pública de João Pessoa – PB. Desse modo, tenta-se explicar o porquê de estudantes com características observáveis semelhantes obterem resultados tão diferenciados. Para alcançar esse objetivo foram utilizados os Microdados da Prova Brasil de 2011. A estratégia empírica adotada consistiu em combinar o método de Pareamento por Escore de Propensão com a Decomposição de Oaxaca-Blinder. Assim, foi possível obter uma amostra mais homogênea em termos de fatores observáveis dos discentes e separá-los em dois grupos: aqueles com desempenho acima e abaixo da média. Criado esses dois grupos, utilizou-se da decomposição de Oaxaca-Blinder para decompor o diferencial de média entre fatores observáveis e não observáveis. O diferencial decorrente de fatores não observáveis foi atribuído ao efeito contingência, ou seja, às características inatas, particulares, de cada estudante. Os resultados obtidos apresentaram evidências favoráveis à ocorrência do referido fenômeno. O segundo ensaio, por sua vez, consiste em um exercício empírico de avaliação de impacto. Buscou-se avaliar os efeitos do Programa Mais Educação, implementado em 2008, sobre cinco variáveis de impacto: o Índice de Desenvolvimento da Educação Básica (IDEB) das escolas, a nota média das escolas em português e em matemática, a taxa de abandono e um índice de violência construído. Optou-se por avaliar esse programa por dois principais motivos: primeiro porque ele vem sendo cada vez mais adotado nas escolas brasileiras, e segundo porque os gastos com o mesmo são consideravelmente elevados. Para tanto, utilizou-se os Microdados da Prova Brasil de 2007, 2009 e 2011. Dessa forma, foi possível construir uma estrutura de dados em painel contendo informações em um período anterior (2007) e em dois períodos posteriores ao início do programa (2009 e 2011). A mensuração dos impactos do Mais Educação deu-se pela combinação de dois métodos amplamente empregados nesse tipo de análise, a saber: Pareamento por Escore de Propensão e Diferenças em Diferenças. A combinação dessas metodologias fornece estimativas robustas do efeito do programa sobre as variáveis de impacto adotadas, pois permite o controle tanto pelas características observáveis como pelas não observáveis. Os modelos estimados indicam que o programa em questão é eficiente, sobretudo, na redução das taxas de abandono e violência nas escolas.

**Palavras-chaves:** Desigualdade de Contingência, Avaliação de Impacto, Mais Educação.

## ABSTRACT

This Master's Thesis consists of two essays that seek to answer two central questions. In the first essay the main goal is to verify the occurrence of contingency inequality in the sense described by Rawls (1971), among students from 5th and 9th grade from public schools of João Pessoa - PB. Thus, it attempts to explain why students with similar observable characteristics get such different results. To achieve this goal, it was used the Microdata from Prova Brasil 2011. The empirical strategy adopted was to combine the Propensity Score Matching with the Oaxaca-Blinder decomposition. This way, it was possible to obtain a more homogeneous sample in terms of observable factors of the students and separate them into two groups: those with performance above and below the average. Created these two groups, the Oaxaca-Blinder decomposition was used to decompose the average differential between observable and unobservable factors. The differential due to unobservable factors was attributed to the contingency effect, ie, the innate characteristics of each student. The results showed evidence favorable to the occurrence of such phenomenon. The second essay consists of an empirical exercise of impact evaluation. We sought to evaluate the effects of the Program Mais Educação, implemented in 2008, over five outcome variables: the Índice de Desenvolvimento da Educação Básica (IDEB) of the schools, the average grade of schools in portuguese and mathematics, the dropout rate and an index of violence that was created. It was decided to evaluate this program for two main reasons: first because it has been increasingly adopted in Brazilian schools, and second because its costs are considerably high. For this purpose, it was used the Microdata from Prova Brasil 2007, 2009 and 2011. Thus, it was possible to construct a panel data structure containing information in a previous period (2007) and in two periods after the beginning of the program (2009 and 2011). The measurement of the impacts of Mais Educação was obtained through the combination of two methods widely used in this type of analysis, namely: Propensity Score Matching and Differences in Differences. The combination of these methods provides robust estimates of the effect of the program on the adopted outcome variables, because it allows the control of both the observable and unobservable characteristics. The estimated models indicate that the program in question is efficient especially in reducing dropout rates and violence in schools.

**Keywords:** Contingency Inequality, Impact Evaluation, Mais Educação.

## LISTA DE FIGURAS

Figura 1 – Função de Densidade do Escore de Propensão (Paraíba - 5º Ano).....	70
Figura 2 – Função de Densidade do Escore de Propensão (Paraíba - 9º Ano).....	70
Figura 3 – Função de Densidade do Escore de Propensão (Nordeste - 5º Ano) .....	72
Figura 4 – Função de Densidade do Escore de Propensão (Nordeste - 9º Ano) .....	72
Figura 5 – Função de Densidade do Escore de Propensão (Brasil - 5º Ano) .....	74
Figura 6 – Função de Densidade do Escore de Propensão (Brasil - 9º Ano) .....	75

## **LISTA DE QUADROS**

Quadro 1 – Descrição das variáveis .....	20
Quadro 2 – Descrição das variáveis .....	47

## LISTA DE TABELAS

Tabela 1 - Estatísticas descritivas dos alunos do 5º e 9º ano do município de João Pessoa ....	23
Tabela 2 – Estimação do escore de propensão (5º Ano) .....	29
Tabela 3 – Estimação do escore de propensão (9º Ano) .....	31
Tabela 4 – Resultados da decomposição de Oaxaca-Blinder .....	33
Tabela 5 – Resultados da decomposição de Oaxaca-Blinder com PSM .....	35
Tabela 6 - Estatísticas descritivas da base de dados completa para os alunos do 5º e 9º ano do município de João Pessoa .....	42
Tabela 7 - Média dos indicadores de impacto para os grupos de controle e tratamento .....	51
Tabela 8 - Estatísticas descritivas dos alunos do 5º e 9º ano da Paraíba, do Nordeste e do Brasil.....	54
Tabela 9 – Estimação dos efeitos do Mais Educação sobre o IDEB .....	62
Tabela 10 – Estimação dos efeitos do Mais Educação sobre a média de português e matemática das escolas .....	63
Tabela 11 – Estimação dos efeitos do Mais Educação sobre a taxa de abandono .....	64
Tabela 12 – Estimação dos efeitos do Mais Educação sobre o índice de violência .....	65
Tabela 13 – Estimação do escore de propensão para as escolas da Paraíba.....	69
Tabela 14 – Estimação do escore de propensão para as escolas do Nordeste.....	71
Tabela 15 – Estimação do escore de propensão para as escolas do Brasil.....	73

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>13</b>
<b>REFERÊNCIAS .....</b>	<b>15</b>
 <b>2. ESTIMAÇÃO DA DESIGUALDADE DE CONTINGÊNCIA ENTRE OS ESTUDANTES DO ENSINO FUNDAMENTAL DA REDE PÚBLICA DE JOÃO PESSOA – PB .....</b>	<b>16</b>
<b>2.1 INTRODUÇÃO .....</b>	<b>16</b>
<b>2.2 FONTE E TRATAMENTO DOS DADOS .....</b>	<b>18</b>
<b>2.3 ESTRATÉGIA DE ESTIMAÇÃO.....</b>	<b>24</b>
2.3.1 PAREAMENTO POR ESCORE DE PROPENSÃO .....	24
2.3.2 DECOMPOSIÇÃO DE OAXACA-BLINDER .....	27
<b>2.4 ANÁLISE E DISCUSSÃO DOS RESULTADOS .....</b>	<b>28</b>
2.4.1 ESTIMAÇÃO DO ESCORE DE PROPENSÃO .....	28
2.4.2 ESTIMAÇÃO DA DECOMPOSIÇÃO DE OAXACA-BLINDER .....	32
2.4.2 ESTIMAÇÃO DA DECOMPOSIÇÃO DE OAXACA-BLINDER COM PSM .....	35
<b>2.5 CONSIDERAÇÕES FINAIS .....</b>	<b>37</b>
<b>REFERÊNCIAS .....</b>	<b>39</b>
<b>APÊNDICE.....</b>	<b>42</b>
 <b>3. AVALIAÇÃO DE IMPACTO DO PROGRAMA MAIS EDUCAÇÃO NAS ESCOLAS DA REDE PÚBLICA DE ENSINO FUNDAMENTAL: UMA ANÁLISE PARA O PERÍODO 2007-2011.....</b>	<b>43</b>
<b>3.1 INTRODUÇÃO .....</b>	<b>43</b>
<b>3.2 PROGRAMA MAIS EDUCAÇÃO .....</b>	<b>45</b>
<b>3.3 FONTE E TRATAMENTO DOS DADOS .....</b>	<b>47</b>
<b>3.4 ESTRATÉGIA EMPÍRICA .....</b>	<b>56</b>
3.4.1 GRUPOS DE TRATAMENTO E DE CONTROLE .....	56
3.4.2 MODELO EMPÍRICO: DIFERENÇAS EM DIFERENÇAS .....	56
3.4.3 ANÁLISE DE COMPONENTES PRINCIPAIS .....	60
<b>3.5 ANÁLISE E DISCUSSÃO DOS RESULTADOS .....</b>	<b>61</b>
3.5.1 IMPACTOS SOBRE O IDEB.....	62
3.5.2 IMPACTOS SOBRE AS NOTAS MÉDIAS DAS ESCOLAS .....	63
3.5.3 IMPACTOS SOBRE A TAXA DE ABANDONO.....	64
3.5.4 IMPACTOS SOBRE O ÍNDICE DE VIOLÊNCIA .....	64
<b>3.6 CONSIDERAÇÕES FINAIS .....</b>	<b>65</b>
<b>REFERÊNCIAS .....</b>	<b>67</b>
<b>APÊNDICE.....</b>	<b>69</b>

## 1. INTRODUÇÃO

A literatura acerca da economia da educação tem avançado consideravelmente desde a publicação dos trabalhos de Gary Becker ao final dos anos 1950 e ao longo das décadas de 1960 e 1970. Apesar de outros pesquisadores terem se debruçado sobre esse tema no mesmo período que Becker, com destaque para o economista americano Theodore Schultz, considerado por muitos o pioneiro nessa temática, foi sobretudo com os trabalhos desse primeiro que a disciplina ganhou corpo e notoriedade na ciência econômica.

Segundo Becker (2011), os estudos da moderna economia da educação tiveram início nos anos 1950, com os trabalhos de Schultz (1963, 1971, 1972, 1982), Mincer (1958, 1974, 1981), Rosen (1979), e dele próprio. Nessa época, a disciplina era vista como sendo parte integrante de um campo mais geral sobre investimentos em capital humano. Com o passar do tempo, todavia, vários economistas, sociólogos e demais interessados passaram a dedicar esforços de modo identificar os mais variados aspectos da formação educacional dos indivíduos. Graças a isso, pode-se dizer que tal disciplina se consolidou e tem atraído cada vez mais a atenção da sociedade.

Inicialmente, o foco dos estudos em economia da educação visavam compreender os ganhos de produtividade oriundos do fator capital humano na produção. Em outras palavras, de que modo uma mão-de-obra mais qualificada impactaria nas atividades produtivas e seria capaz de explicar os diferenciais de crescimento e desenvolvimento econômico entre as nações. Schultz (1963) considerava a educação como o investimento básico em capital humano e cuja contribuição seria multidimensional, servindo a fins sociais, políticos, etc. O autor faz uma profunda análise sobre os retornos e custos da educação e conclui que a instrução é o maior investimento que homem pode proporcionar a si mesmo, uma vez que isso o tornará um trabalhador mais produtivo e capaz de obter melhores retornos no mercado de trabalho.

Já em meados da década de 1970, a agenda de pesquisa relacionada à área da Economia da Educação se dedicou a estudar os determinantes do desempenho escolar e, para isso, foram consideradas características observáveis do aluno, o *background* familiar, a estrutura física e a gestão escolar, bem como aspectos econômicos e demográficos dos professores e diretores. Aliado a isto, também se observa um crescente interesse na avaliação do desempenho a partir de testes e exames padronizados.

No entanto, alguns aspectos antes negligenciados nas pesquisas teóricas e empíricas devido à dificuldade de mensuração, como habilidades socioemocionais, passaram a ser

estudadas a partir do trabalho de Heckman, Stixrud e Urzua (2006). Diante disso, pode-se dizer que a função de produção da educação contempla uma gama de características.

Tendo em vista os recentes avanços que têm sido realizados nessa literatura, a presente dissertação busca por meio de dois ensaios analisar o desempenho escolar dos alunos do Ensino Fundamental.

O primeiro ensaio estuda o papel das características não observáveis dos alunos, buscando atribuir um novo significado ao efeito dessas características, de modo a explicar melhor o papel desempenhado pelas mesmas e sua importância no desempenho dos estudantes. Já no segundo ensaio, o foco é sobre os impactos do Programa Federal Mais Educação sobre alguns indicadores escolares como, proficiência, abandono e violência escolar. Para isso, utiliza-se de técnicas de avaliação de políticas públicas para tentar aferir o efeito causal do referido programa.

Em suma, pode-se dizer que o primeiro ensaio busca contribuir de forma teórica e empírica, enquanto que o segundo consiste em um exercício de avaliação de impacto. Diante disso, ao abordar temas educacionais considerando características observáveis e não observáveis dos alunos da rede pública de ensino, aspectos do ambiente que os influencia, bem como os efeitos de uma política pública adotada em grande parte das escolas brasileiras, espera-se que as informações fornecidas por esta dissertação possam servir de base para futuros estudos e formulação de políticas educacionais.



## REFERÊNCIAS

BECKER, G. S. Reflections on the economics of education. In: HANUSHEK, E. A.; MACHIN, S.; WOESSMANN, L. (Orgs.). **Handbook of the Economics of Education (Book 4)**. 1. ed. United States of America: North Holland, 2011. p. 11-14.

HECKMAN, J. J.; STIXRUD, J.; URZUA, S. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. **Journal of Labor Economics**, University of Chicago Press, v.. 24(3), p. 411-482, 2006.

MINCER, J. Investment in Human Capital and Personal Income Distribution. **Journal of Political Economy**, v. 66, n. 4, p. 281-302, 1958.

MINCER, J. **Schooling, Experience, and Earnings**. New York: Columbia University Press, 1974.

MINCER, J. Human Capital and Economic Growth. **Working Paper n. 803**. National Bureau of Economic Research, 1981.

ROSEN, S.; WILLIS, R. J. Education and self-selection. **The Journal of Political Economy**, v. 87, n. 5, p. S7-S36, 1979.

SCHULTZ, T. W. **The Economic Value of Education**. New York: Columbia University Press, 1963.

SCHULTZ, T. W. **Investment in Human Capital: The Role of Education and of Research**. New York: Free Press, 1971.

SCHULTZ, T. W. Human Capital: Policy Issues and Research Oportunities. In: **Economic Research: Retrospect and Prospect**. Vol 6: Human Resources. National Bureau of Economic Research, 1972. p. 1-84.

SCHULTZ, T. W. **Investing in People: The Economics of Population Quality**. London: University of California Press, 1982.

## **2. ESTIMAÇÃO DA DESIGUALDADE DE CONTINGÊNCIA ENTRE OS ESTUDANTES DO ENSINO FUNDAMENTAL DA REDE PÚBLICA DE JOÃO PESSOA – PB**

### **2.1 INTRODUÇÃO**

A questão dos determinantes do desempenho educacional tem levantado discussões nas diferentes esferas do governo, entidades de classe e na sociedade como um todo. Conforme apontado por Sen (1985), o desempenho educacional é fator determinante tanto para a formação profissional do indivíduo como para seu próprio bem-estar. O desempenho escolar também costuma ser correlacionado com desigualdade de renda e divergências no *status* de saúde dos indivíduos, tornando-se assim fator relevante nas decisões políticas das nações desenvolvidas e subdesenvolvidas (HESHMATI, 2004).

A partir da década de 1970, diferentes pesquisas e testes padronizados<sup>1</sup> têm objetivado mensurar o desempenho educacional e cognitivo dos estudantes em diversos países, levando a comparações intra e entre nações no que se reporta à dispersão das notas dos estudantes. Os resultados desses testes têm motivado as diferentes sociedades a identificar os determinantes do desempenho escolar, sendo que cada vez mais as características do ambiente familiar, condições da escola, incentivo ao estudo dado por familiares e/ou pelo professor, bem como o esforço dos alunos e outras características individuais vêm sendo apontados como fatores que contribuem largamente para a melhoria do aprendizado.

Adicionalmente, uma literatura mais recente investiga o papel das questões de justiça e equidade nos resultados educacionais. A partir da tríade justiça, desigualdade e desempenho, o papel da desigualdade de oportunidade vem sendo adotado como um dos preditores da acumulação de capital humano (PATTANAIK e XU, 1990; KRANICH, 1996; THOMSON, 2011; FIGUEIREDO e SILVA, 2011). Contudo, não há um consenso sobre como seria possível promover uma disseminação da educação de modo mais justo.

Considerando o conceito de Rawls (1971), existirá justiça se todos os indivíduos tiverem acesso às mesmas condições e benefícios, o que tende a garantir a equidade. Entretanto, a teoria rawlsiana considera ainda que aos indivíduos diferentes sejam dados tratamentos diferenciados. Dessa forma, essa teoria considera que alguns membros da sociedade apresentam melhores dotações como, por exemplo, maior habilidade e/ou talento nato. Sendo assim, estes indivíduos

---

<sup>1</sup> Com destaque para o *Programme for International Student Assessment* (PISA).

são mais privilegiados *vis-à-vis* outros sujeitos da sociedade, de tal modo que ao segundo grupo deveria ser dada atenção especial em decorrência da desigualdade de contingência.

A desigualdade de contingência representa, portanto, a diferença no conjunto de habilidades inatas que alguns indivíduos possuem e que interfere no processo de aprendizagem. Desse modo, mesmo em condições de plena igualdade de oportunidade é de se esperar que os resultados dos indivíduos sejam diferentes em decorrência das particularidades que cada um possui<sup>2</sup>.

Com relação ao papel da educação na ótica de Rawls, Gondim (2010) e Gondim e Rodrigues (2011) desenvolvem discussões mais detalhadas. Segundo esses autores, para Rawls, a educação poderia ser definida como o treinamento e o desenvolvimento de habilidades e aptidões capazes de conceder maior autonomia aos cidadãos, tornando-os mais preparados para o debate público e favorecendo o bom funcionamento de uma sociedade. Além disso, uma vez que os indivíduos são diferentes em suas dotações, caberia à sociedade alocar recursos à educação de modo a corrigir “os desvios de contingência na direção da igualdade”<sup>3</sup>.

Em vista do que foi abordado, a motivação da presente pesquisa parte de duas frentes. Do ponto de vista teórico, testar-se-á a hipótese de que os indivíduos diferentes, mesmo que submetidos às mesmas condições iniciais e ao mesmo tratamento, ainda assim apresentarão resultados diferentes, tanto do ponto de vista da construção dos níveis de capital humano como na capacidade de edificação das habilidades. Nestes termos, pode-se dizer que esse estudo se propõe a investigar o desempenho escolar a partir das notas de proficiência de português e matemática para os alunos do 5º e 9º ano das escolas públicas de João Pessoa, na Paraíba, aferidos a partir da Prova Brasil de 2011.

Visando uma maior confiabilidade nos resultados, foi utilizado o método de Pareamento por Escore de Propensão (*Propensity Score Matching* - PSM) para controlar o desempenho escolar dos alunos por características observáveis, tais como estrutura física e organizacional da escola, nível de instrução dos professores e diretores, esforço dos alunos, incentivo da família aos estudos, entre outros.

Espera-se que mesmo considerando condições iguais – sejam elas familiares ou de oportunidade, esforço e incentivo – o desempenho nos exames padronizados seja divergente em função da diferença de habilidades, o que não garante uma distribuição justa de contingência, haja vista que os indivíduos mais habilidosos são comparativamente mais

---

<sup>2</sup> Hanushek (1970) apresenta uma função de produção para a educação que inclui um vetor de características inatas como insumo para o desempenho escolar.

<sup>3</sup> Rawls (1971), p. 101.

favorecidos. Com isso, muitas dimensões das desigualdades no rendimento escolar dos diferentes estudantes não podem ser captadas apenas por fatores observáveis, dado que os indivíduos são dissimilares em sua própria natureza, caracterizando assim o efeito contingência.

Em suma, a partir de uma perspectiva prática, a hipótese é de que mesmo controlando por fatores observáveis, o desempenho em português e matemática dos alunos tanto do 5º quanto do 9º ano são díspares em relação aos indivíduos semelhantes. Portanto, admite-se que a diferença em termos de desempenho seja decorrente da divergência inata de cada indivíduo.

A estratégia empírica adotada para alcançar o objetivo proposto consiste em estimar um modelo por meio da técnica de PSM para gerar um suporte comum, o que tornará a amostra mais homogênea em termos de características observáveis. A partir disso, será possível comparar o desempenho dos alunos. Em seguida, emprega-se a metodologia desenvolvida por Oaxaca e Blinder (1973) para decompor o desempenho estudantil de português e de matemática entre componentes observáveis (*background* familiar, estrutura da escola, características do aluno, do professor e diretor) e não observáveis (desigualdade de contingência, no sentido descrito por Rawls)<sup>4</sup>. A expectativa de diferenças significativas para os componentes não observados fornece evidências de que existe efeito contingência entre os alunos do 5º e 9º ano das escolas públicas de João Pessoa.

Além dessa introdução, essa pesquisa contém mais quatro seções. A próxima seção apresenta a fonte e o tratamento dos dados utilizados. Em seguida, discute-se em detalhes a estratégia de estimação empregada, sendo acompanhada logo após pela análise dos resultados alcançados. Finalmente, a última seção contempla as considerações finais.

## 2.2 FONTE E TRATAMENTO DOS DADOS

Para a realização desta pesquisa, utilizou-se os microdados da Prova Brasil de 2011, disponíveis no portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A Prova Brasil é uma das três modalidades que compõe o Sistema de Avaliação da Educação Básica (SAEB), implementado pelo Governo Federal em 1990. Juntamente com a Prova Brasil, fazem parte do SAEB a Avaliação Nacional da Educação Básica (ANEB) e a Avaliação Nacional da Alfabetização (ANA).

---

<sup>4</sup> É importante ressaltar que a decomposição de Oaxaca-Blinder foi inicialmente proposta para captar diferencial de salários, atribuindo às características não observáveis os efeitos da discriminação. O que se fez nesse estudo foi empregar este procedimento para decompor o diferencial de nota entre estudantes e dar um novo sentido a estas características não observáveis.

A ANEB consiste na parte amostral do SAEB, abrangendo alunos das redes públicas e privadas do país, tanto em áreas urbanas quanto em áreas rurais, matriculados no 5º e 9º ano do Ensino Fundamental e no 3º ano do Ensino Médio. O público-alvo dessa modalidade é composto por escolas que tenham de 10 a 19 estudantes matriculados no 5º ou 9º ano do Ensino Fundamental regular público, escolas que tenham 10 ou mais estudantes matriculados no 5º ou 9º ano do Ensino Fundamental regular privado, e escolas que tenham 10 ou mais estudantes matriculados no 3º ano do Ensino Médio regular público ou privado (BRASIL, 2012).

A Avaliação Nacional do Rendimento Escolar (ANRESC), comumente chamada de Prova Brasil, é uma avaliação censitária, envolvendo alunos do 5º e 9º ano do Ensino Fundamental das escolas públicas das redes municipais, estaduais e federal. São consideradas todas as escolas com pelo menos 20 alunos matriculados, nas áreas rurais e urbanas do Brasil (BRASIL, 2012).

Por fim, tem-se a ANA<sup>5</sup>, que se trata também de uma avaliação censitária, englobando os alunos do 3º ano do Ensino Fundamental das escolas públicas. Seu objetivo principal é avaliar os níveis de alfabetização e letramento em Língua Portuguesa, alfabetização Matemática e condições de oferta do Ciclo de Alfabetização das redes públicas (BRASIL, 2012).

Isto posto, como o objetivo dessa pesquisa é captar o possível diferencial de desempenho existente entre os estudantes do Ensino Fundamental da rede pública de João Pessoa, controlando as comparações por meio de atributos observáveis e usando o método de Oaxaca-Blinder para decompor essa diferença em fatores explicados e não explicados, tem-se que os microdados da Prova Brasil são adequados à essa finalidade. Isso ocorre, pois, seus questionários apresentam informações referentes às características dos discentes e seus domicílios, além de fornecer informações importantes acerca das escolas, do diretor e dos professores. Dessa forma, esta se configura em uma base de dados que permite comparar estudantes que possuem características similares, voltada exclusivamente à rede pública do Ensino Fundamental.

No total os microdados da Prova Brasil de 2011 contém informações sobre 5.201.730 estudantes brasileiros. Porém, conforme estipulado nos objetivos e a fim de tornar a base mais homogênea, considerou-se apenas os discentes do município de João Pessoa. Após esse recorte e a exclusão das observações faltantes (*missing values*), a base final dispõe de 1.193 observações para os alunos do 5º ano, e 2.486 observações para os alunos do 9º ano.

---

<sup>5</sup> A Avaliação Nacional da Alfabetização somente foi incorporada ao SAEB em 2013. Tal avaliação ocorre todos os anos, enquanto que a ANEB e a ANRESC são realizadas bianualmente.

Como o desempenho escolar pode ser afetado por características dos alunos e do domicílio, bem como pelo ambiente familiar e situação socioeconômica, faz-se necessário controlar por essas características para aferir o efeito contingência dos alunos. Sendo assim, optou-se por reportar no Quadro 1 a descrição das variáveis utilizadas, ressaltando que a escolha por controlar por determinadas características é feita a partir da literatura acerca dos determinantes do desempenho escolar (BARROS et al., 2001; FELÍCIO e FERNANDES, 2005; SOUZA, 2005; CURI e MENEZES-FILHO 2006; MENEZES-FILHO, 2007; MACHADO et al., 2008 ARAÚJO e ALMEIDA, 2013; VERNIER e BAGOLIN, 2013; ALMEIDA, 2014).

As variáveis referentes à leitura de livros e resposta da tarefa de casa são utilizadas como *proxies* para medir esforço e dedicação, respectivamente. Com relação às variáveis dos demais grupos, alusivas às condições das escolas, escolaridade e experiência dos professores e diretores, estas têm como objetivo medir os efeitos do ambiente escolar como um todo sobre o desempenho dos estudantes.

**Quadro 1 – Descrição das variáveis**

<b>Variáveis</b>	<b>Descrição das Variáveis</b>
Proficiência_SAEB_português	Proficiência do aluno em língua portuguesa calculada na escala única do SAEB.
Proficiência_SAEB_matemática	Proficiência do aluno em matemática calculada na escala única do SAEB.
<b>Características dos Alunos e do Domicílio</b>	
Sexo_aluno_masculino	1 se o sexo do aluno for masculino e 0 caso contrário.
Raça_aluno_branco	1 se a raça do aluno for branca e 0 caso contrário.
Trabalho_aluno	1 se o aluno trabalha fora de casa e 0 caso contrário.
Incentivo_aluno	1 se os pais do aluno o incentivam a estudar e 0 caso contrário.
Ler_livros_sempre	1 se o aluno ler livros sempre ou quase sempre e 0 caso contrário.
Faz_tarefa_sempre <sup>6</sup>	1 se o aluno faz a tarefa de casa sempre ou quase sempre e 0 caso contrário.
Creche_ou_pré-escola	1 se o aluno iniciou os estudos na creche ou pré-escola e 0 caso contrário.
Reprovação_aluno	1 se o aluno nunca foi reprovado e 0 caso contrário.
Computador_com_internet	1 se na casa do aluno tem computador com internet e 0 caso contrário.
Escol_mãe_ens_superior	1 se a mãe do aluno completou a faculdade e 0 caso contrário.

<sup>6</sup> Ao avaliar a proficiência em português (matemática), a variável considerada foi se o estudante faz a tarefa de casa de português (matemática) sempre ou quase sempre.

Prof_corrige_tarefa_sempre <sup>7</sup>	1 se o professor corrige a tarefa de casa sempre ou quase sempre e 0 caso contrário.
<b>Características das Escolas<sup>8</sup></b>	
Salas_escola_bom	1 se o estado de conservação das salas de aula da escola é considerado bom e 0 caso contrário.
Banheiros_escola_bom	1 se o estado de conservação dos banheiros da escola é considerado bom e 0 caso contrário.
Biblioteca_escola_bom	1 se o estado de conservação da biblioteca da escola é considerado bom e 0 caso contrário
Laboratório_escola_bom	1 se o estado de conservação do laboratório da escola é considerado bom e 0 caso contrário
<b>Características dos Professores</b>	
Escol_professor_ens_superior	1 se o professor concluiu o ensino superior e 0 caso contrário.
Experiência_professor	1 se o professor leciona há 10 anos ou mais e 0 caso contrário.
<b>Características do Diretor</b>	
Escol_diretor_ens_superior	1 se o diretor possui ensino superior completo e 0 caso contrário
Experiência_diretor	1 se o diretor trabalha em educação há 10 anos ou mais e 0 caso contrário

Fonte: Microdados da Prova Brasil de 2011.

Apresentada a fonte dos dados e a descrição das variáveis, a Tabela 1 contempla as estatísticas descritivas da amostra<sup>9</sup>. Como pode ser observado, cerca de 49,29% dos alunos do 5º ano são do sexo masculino, enquanto que no 9º ano tal proporção cai para 45,10%<sup>10</sup>. Além disso, 25,80% e 24,29% dos discentes do 5º e 9º ano, respectivamente, declararam ser da cor branca. Em ambas as séries, mais de 11% dos alunos declararam possuir algum tipo de trabalho fora de casa, e mais de 97% afirmaram que seus pais os incentivam a estudar.

Algumas diferenças consideráveis podem ser apontadas nas variáveis de leitura de livros e de resposta da tarefa de casa. Enquanto que mais de 45% dos alunos do 5º ano alegaram ler livros sempre ou quase sempre, apenas 28,10% dos estudantes do 9º afirmaram possuir tal hábito. Já no que diz respeito às tarefas de casa, quase 80% dos discentes do 5º ano afirmaram respondê-las com frequência, contra menos de 56% dos alunos do 9º ano.

Com relação às características das escolas, em geral, as principais diferenças residem nas condições da biblioteca e do laboratório. A proporção de escolas com biblioteca em boas

<sup>7</sup> Ao avaliar a proficiência em português (matemática), a variável considerada foi se o professor corrige a tarefa de casa de português (matemática) sempre ou quase sempre.

<sup>8</sup> Todas as escolas públicas de Ensino Fundamental do município de João Pessoa, contempladas pela Prova Brasil 2011, possuem dependência administrativa municipal ou estadual.

<sup>9</sup> As estatísticas descritivas da base de dados completa, isto é, sem a exclusão das observações faltantes (*missing values*) podem ser consultadas no apêndice.

<sup>10</sup> Comparando com os dados da PNAD 2011, tem-se que 51,25% dos indivíduos paraibanos entre 9 e 11 anos são do sexo masculino, enquanto que 48,75% são do sexo feminino. Considerando esta mesma informação para a faixa etária entre 13 e 15 anos, a relação cai para 50,36% de homens e 49,64% de mulheres.

condições é maior para os alunos do 9º ano. Já com relação às condições do laboratório, a proporção é maior para os alunos do 5º ano. Vale lembrar que algumas escolas ofertam tanto o 5º quanto o 9º ano, e todas estão localizadas no mesmo município, isto é, João Pessoa.

No que se refere aos professores, constatou-se que mais de 99% dos que ensinam para os alunos do 9º ano concluíram o ensino superior, contra 87% dos professores dos alunos do 5º ano. No quesito experiência, todavia, tem-se que, na média, os professores do 5º ano são mais experientes. Por fim, tem-se que mais de 94% dos diretores das duas séries concluíram o ensino superior e trabalham na área da educação há pelo menos 10 anos.



**Tabela 1 - Estatísticas descritivas dos alunos do 5º e 9º ano do município de João Pessoa**

	Variáveis	5º ano				9º ano			
		Média	Desvio-padrão	Mínimo	Máximo	Média	Desvio-padrão	Mínimo	Máximo
<b>Características do Aluno e do Domicílio</b>	Sexo_aluno_masculino	0,4929	0,5001	0	1	0,4510	0,4977	0	1
	Raça_aluno_branco	0,2580	0,4377	0	1	0,2429	0,4289	0	1
	Trabalho_aluno	0,1158	0,3200	0	1	0,1113	0,3146	0	1
	Incentivo_aluno	0,9724	0,1639	0	1	0,9861	0,1170	0	1
	Ler_livros_sempre	0,4547	0,4981	0	1	0,2810	0,4496	0	1
	Faz_tarefa_português_sempre	0,7919	0,4061	0	1	0,5533	0,4972	0	1
	Faz_tarefa_matemática_sempre	0,7815	0,4134	0	1	0,5152	0,4998	0	1
	Creche_ou_pré-escola	0,7882	0,4087	0	1	0,8527	0,3545	0	1
	Reprovação_aluno	0,6397	0,4802	0	1	0,5668	0,4956	0	1
	Computador_com_internet	0,3800	0,4856	0	1	0,5089	0,5000	0	1
	Escol_mãe_ens_superior	0,1730	0,3784	0	1	0,0966	0,2955	0	1
	Prof_corrige_tarefa_port_sempre	0,7800	0,4144	0	1	0,7489	0,4337	0	1
	Prof_corrige_tarefa_mat_sempre	0,7933	0,4051	0	1	0,7466	0,4351	0	1
	Proficiência_português	191,0498	42,9642	90,9695	320,6530	244,5807	44,15711	108,8913	360,7970
	Proficiência_matemática	206,9611	42,1496	99,7824	337,1514	245,3179	44,81949	120,9378	396,0480
<b>Características da Escola</b>	Salas_escola_bom	0,7282	0,4450	0	1	0,7570	0,4290	0	1
	Banheiros_escola_bom	0,5039	0,5002	0	1	0,4258	0,4945	0	1
	Biblioteca_escola_bom	0,6940	0,4610	0	1	0,8185	0,3855	0	1
	Laboratório_escola_bom	0,5347	0,4990	0	1	0,4448	0,4970	0	1
<b>Características do Professor</b>	Escol_professor_ens_superior	0,8693	0,3372	0	1	0,9948	0,0721	0	1
	Experiência_professor	0,7204	0,4489	0	1	0,6159	0,4865	0	1
<b>Características do Diretor</b>	Escol_diretor_ens_superior	0,9472	0,2237	0	1	0,9798	0,1406	0	1
	Experiência_diretor	0,9334	0,2494	0	1	0,9860	0,1175	0	1

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Nota: Valores expandidos pelo peso amostral.

## 2.3 ESTRATÉGIA DE ESTIMAÇÃO

Para a consecução dos objetivos delineados, primeiramente empregou-se o método de Pareamento por Escore de Propensão. Feito isso, utilizou-se da decomposição de Oaxaca-Blinder para aferir a desigualdade de contingência entre os alunos do 5º e 9º ano. Neste sentido, inicialmente, a amostra foi dividida entre os estudantes do 5º e do 9º ano do Ensino Fundamental da rede pública de João Pessoa e, a partir das notas de proficiência em português e matemática, estes estudantes foram divididos em dois grupos: aqueles com desempenho acima da média e aqueles com desempenho abaixo<sup>11</sup>.

Uma vez divididos os grupos para cada ano, utilizou-se variáveis referentes às características dos alunos, dos domicílios, das escolas, dos professores e dos diretores para garantir que cada estudante tenha seu desempenho comparado apenas com outros estudantes que sejam similares nos atributos selecionados. Caso seja verificado que há diferença estatisticamente significativa no desempenho médio dos dois grupos, é possível utilizar o procedimento de Oaxaca-Blinder para decompor essa diferença em fatores explicados e não explicados. Os fatores não explicados (não observáveis) será atribuído ao efeito contingência, isto é, o conjunto de habilidades inatas que alguns indivíduos possuem e que interfere no processo de aprendizado.

Com base nessa descrição, pode-se dizer que a estratégia de estimação do efeito contingência se dá por meio da combinação do método de Pareamento por Escore de Propensão seguido pela decomposição de Oaxaca-Blinder<sup>12</sup>. Diante disso, optou-se por apresentar em mais detalhes cada um desses métodos.

### 2.3.1 PAREAMENTO POR ESCORE DE PROPENSÃO

O método de pareamento tem como objetivo reduzir o viés de autosseleção presente na estimação do impacto de um dado tratamento. Tal método busca construir um grupo de controle consistente e o mais similar possível ao grupo de tratamento, partindo da concepção de que quanto mais parecido forem esses grupos em suas características observáveis, menor será o viés.

---

<sup>11</sup> As médias de proficiência em português e matemática, tanto para o 5º quanto para o 9º ano, podem ser consultadas na Tabela 1.

<sup>12</sup> Ao combinar esses dois procedimentos é possível corrigir o viés apontado por Ñopo (2008), que critica a decomposição de Oaxaca-Blinder por considerar na estimação as observações que estariam dentro e fora da região de suporte comum. Segundo esse autor o correto seria considerar apenas as observações que estariam dentro dessa região.

Como ressaltado em Menezes Filho (2012, p. 85): “de acordo com as hipóteses desse método, cada membro do grupo de tratamento teria um par no grupo de controle que representa o resultado que ele teria obtido caso não fosse tratado”. Assim, o pareamento assume que ao controlar os grupos com base em um conjunto de atributos observáveis, o grupo de controle passaria a representar um contrafactual adequado ao grupo de tratamento.

Apesar do referido método ser usado sobretudo para avaliar os efeitos de políticas públicas e programas sociais, pode-se empregá-lo basicamente em qualquer situação onde seja possível separar as unidades de observação em grupos com base em uma dada variável de resultado. Feito essa divisão, o método consiste em comparar (parear) essas unidades usando como critério o conjunto de características observáveis pré-existentes selecionadas pelo pesquisador.

É importante ressaltar, todavia, que quanto maior for número de características selecionadas, isto é, quanto maior a dimensão do vetor  $X$ , mais difícil acaba sendo parear as observações (MENEZES FILHO, 2012)<sup>13</sup>. Visando solucionar essa dificuldade, Rosenbaum e Rubin (1983) propuseram o método de *Propensity Score Matching* (PSM), que consiste em utilizar uma função de  $X$  que resuma todas as características presentes nesse vetor.

Isto posto, o escore de propensão pode ser entendido como a probabilidade de receber o tratamento condicionada às variáveis de controle incluídas em  $X$ . Mais formalmente, segundo Becker e Ichino (2002), o escore de propensão é dado por:

$$p(X) \equiv \Pr(D = 1|X) \quad (1)$$

Em que  $D = \{0,1\}$  indica a exposição ao tratamento e  $X$  é o vetor de características observáveis.

Na prática, a estimação do escore de propensão é geralmente feita por meio de procedimentos paramétricos, tais como os modelos *logit* ou *probit* (MENEZES FILHO, 2012). No caso do modelo *logit*, que será o procedimento adotado nesse trabalho, o escore de propensão será estimado como:

$$\hat{p}(x) = \frac{\exp(x\hat{\beta})}{1 + \exp(x\hat{\beta})} \quad (2)$$

---

<sup>13</sup> Tal problema ocorre, sobretudo, se as variáveis forem contínuas.

Em que  $\hat{\beta}$  é o estimador do parâmetro  $\beta$  e  $x$  é o vetor contendo as variáveis de controle selecionadas.

Conforme Khandker, Koolwal e Samad (2010), a validade do PSM necessita que duas hipóteses sejam atendidas: independência condicional e presença de suporte comum. A hipótese de independência condicional, também chamada de não-confundimento, postula que apenas os fatores observáveis afetam a adesão ao programa. Com isso:

$$Y_1, Y_0 \perp D | X \quad (3)$$

Como ressaltado pelos autores, essa é uma hipótese forte que não pode ser diretamente testada, uma vez que depende das propriedades específicas e do desenho de cada programa.

A segunda hipótese, presença de suporte comum, requer que haja sobreposição nos escores de propensão entre as unidades dos grupos de tratado e controle (KHANDKER, KOOLWAL e SAMAD, 2010). De modo mais formal, é preciso que:

$$0 < \Pr[D = 1 | X] < 1 \quad (4)$$

Essa hipótese garante que as unidades no grupo de tratamento possuem um par no grupo de controle que permita sua comparação (MENEZES FILHO, 2012).

Para empregar o referido método nessa pesquisa, será criada uma variável dicotômica que assume valor 1 caso o resultado do aluno na Prova Brasil tenha sido acima da média dos estudantes de João Pessoa, e 0 caso contrário<sup>14</sup>. Tal variável dicotômica será regredida através de um modelo *logit*, tendo como variáveis explicativas as que foram reportadas no Quadro 1. Após a estimação do escore de propensão obtém-se uma região de suporte comum que permitirá comparar estudantes semelhantes em seus atributos observáveis. Uma vez concluída essa primeira etapa, isto é, estimado o escore de propensão e pareado a amostra, pode-se partir para a decomposição de Oaxaca-Blinder de modo a mensurar as parcelas do diferencial total de desempenho dos alunos (em português e matemática) que pode ser atribuído aos fatores observáveis e não observáveis.

---

<sup>14</sup> Mais especificamente, será criada uma variável dicotômica desse tipo tanto para português quanto para matemática, para o 5º e 9º ano.

### 2.3.2 DECOMPOSIÇÃO DE OAXACA-BLINDER

A técnica de decomposição contrafactual desenvolvida por Oaxaca (1973) e Blinder (1973) tem sido vastamente empregada para estudar diferenças de média entre grupos, em relação a uma dada variável de resultado. Essa decomposição se deve aos fatores explicados (observáveis) e aos fatores não explicados (não observáveis), sendo muitas vezes empregada como forma de mensurar discriminação (JANN, 2008).

Para entender como funciona tal procedimento, suponha dois grupos ( $A$  e  $B$ ) e uma variável de resultado ( $Y$ ). A diferença de média de  $Y$  entres esses dois grupos ( $R$ ) pode ser calculada da seguinte forma:

$$R = E(Y_A) - E(Y_B) \quad (5)$$

Em que  $E(Y_A)$  e  $E(Y_B)$  denotam, respectivamente, o valor esperado da variável de resultado para os grupos  $A$  e  $B$ . Considere agora um modelo de regressão linear do tipo:

$$Y_i = X_i' \beta_i + \varepsilon_i \quad (6)$$

Onde  $X$  é um vetor de características determinantes da variável de resultado,  $\beta$  um vetor de parâmetros desconhecidos a ser estimado,  $\varepsilon$  o termo de erro estocástico,  $i \in (A, B)$  e, por hipótese, que  $E(\varepsilon_i) = 0$ . Substituindo (6) em (5) e simplificando os termos:

$$R = E(Y_A) - E(Y_B) = E(X_A)' \beta_A - E(X_B)' \beta_B \quad (7)$$

Assuma agora que existe um vetor de coeficientes não discriminatórios ( $\beta^*$ ), responsável por determinar a contribuição das diferenças entre os regressores. Caso ocorra discriminação contra o grupo  $B$ , mas não haja discriminação em favor do grupo  $A$ , então  $\beta^* = \beta_A$  e a decomposição de Oaxaca-Blinder consiste em estimar:

$$\hat{R} = (\bar{X}_A - \bar{X}_B)' \hat{\beta}_A + \bar{X}_B' (\hat{\beta}_A - \hat{\beta}_B) \quad (8)$$

Em que  $\bar{X}_A$  e  $\bar{X}_B$  representam os valores médios das características dos grupos  $A$  e  $B$ , respectivamente, sendo usados como estimativas para  $E(X_A)$  e  $E(X_B)$ , e  $\hat{\beta}_A$  e  $\hat{\beta}_B$  são os parâmetros estimados por Mínimos Quadrados Ordinários de  $\beta_A$  e  $\beta_B$ <sup>15</sup>.

---

<sup>15</sup> Da mesma forma, caso exista discriminação em favor do grupo  $A$ , mas sem discriminação contra o grupo  $B$ , tem-se que  $\beta^* = \beta_B$  e a estimação será dada por:  $\hat{R} = (\bar{X}_A - \bar{X}_B)' \hat{\beta}_B + \bar{X}_A' (\hat{\beta}_A - \hat{\beta}_B)$ .

A primeira parte da equação (8), ou seja,  $(\bar{X}_A - \bar{X}_B)' \hat{\beta}_A$ , corresponde à parcela do diferencial da variável de resultado que é explicada pelas diferenças observáveis nas características dos grupos. Nesse componente, considera-se que ambos os grupos são valorados da mesma forma e, caso seu valor seja positivo, então os indivíduos do grupo A possuem maiores valores de  $Y$  devido aos seus atributos mensuráveis.

Já a segunda parte, isto é,  $\bar{X}_B'(\hat{\beta}_A - \hat{\beta}_B)$ , capta o potenciais efeitos das características não observáveis, geralmente atribuído à discriminação.<sup>16</sup> Nessa parcela, assume-se que os indivíduos do grupo B apresentam as mesmas características observáveis médias do grupo A. Logo, caso esse componente seja positivo, então os indivíduos do grupo A recebem maiores valores de  $Y$  devido aos seus atributos não observáveis, ou devido a uma possível discriminação contra os indivíduos do grupo B.

Nesse trabalho, a variável de resultado utilizada será a proficiência em português e matemática dos alunos do 5º e 9º ano. Já as variáveis explicativas serão selecionadas com base nos resultados do escore de propensão. Tendo em vista como a decomposição de Oaxaca-Blinder funciona, esta será utilizada para captar o diferencial de desempenho entre os estudantes, examinando a ocorrência de desigualdade de contingência entre alunos com características observáveis semelhantes.

## 2.4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Esta seção contempla a análise e discussão dos resultados que tem como objetivo aferir se existe desigualdade de contingência entre os alunos do 5º e 9º ano da rede pública de João Pessoa, considerando as notas de proficiência em português e matemática. Para tanto, inicialmente, será calculado o escore de propensão para cada série e proficiência, o que permitirá identificar os fatores determinantes do desempenho dos alunos. Além disso, através desse escore, também será possível parear a amostra, tornando-a mais homogênea. Em seguida, são reportados os resultados da decomposição de Oaxaca-Blinder com e sem pareamento, para testar a hipótese de desigualdade de contingência.

### 2.4.1 ESTIMAÇÃO DO ESCORE DE PROPENSÃO

Na Tabela 2 encontram-se os coeficientes estimados para o cálculo do escore de propensão dos alunos do 5º ano. O referido método foi empregado com o intuito de estimar a

---

<sup>16</sup> Nesse trabalho, tais efeitos serão abordados como efeito da desigualdade de contingência no sentido de Rawls.

probabilidade do estudante obter desempenho acima da média na Prova Brasil de 2011. Como ressaltado na seção da metodologia, tal procedimento foi realizado através de um modelo de regressão *logit*, cujas variáveis explicativas são aquelas definidas no Quadro 1.

**Tabela 2 – Estimação do escore de propensão (5º Ano)**

Variáveis Explicativas	Português			Matemática		
	Coefficiente	Erro-padrão	dy/dx	Coefficiente	Erro-padrão	dy/dx
Sexo_aluno_masculino	-0,5523***	0,1136	-0,1371	0,2710**	0,1133	0,0675
Raça_aluno_branco	0,0171	0,1276	0,0043	-0,0697	0,1266	-0,0174
Computador_com_internet	0,1600	0,1187	0,0399	0,3435***	0,1181	0,0853
Escol_mãe_ens_superior	0,0125	0,1524	0,0031	0,0238	0,1512	0,0059
Incentivo_aluno	0,8510**	0,3658	0,2040	2,0786***	0,5019	0,4095
Ler_livros_sempre	-0,2007*	0,1136	-0,0501	-0,2855**	0,1122	-0,0712
Trabalho_aluno	-0,6044***	0,1815	-0,1489	-0,2897*	0,1756	-0,0723
Creche_ou_pré-escola	0,3674***	0,1382	0,0915	0,3984***	0,1367	0,0992
Reprovação_aluno	0,7094***	0,1180	0,1755	0,6955***	0,1175	0,1721
Faz_tarefa_sempre	0,3363**	0,1442	0,0839	0,1115	0,1435	0,0279
Prof_corrige_tarefa_sempre	-0,1015	0,1391	-0,0253	-0,0815	0,1464	-0,0203
Salas_escola_bom	0,0619	0,1453	0,0155	0,2480	0,1647	0,0619
Banheiros_escola_bom	—	—	—	-0,1940	0,1449	-0,0484
Biblioteca_escola_bom	0,0623	0,1493	0,0156	0,3876***	0,1467	0,0966
Laboratório_escola_bom	-0,0104	0,1399	-0,0026	0,0072	0,1422	0,0018
Escol_diretor_ens_superior	0,1266	0,2765	0,0316	—	—	—
Experiência_diretor	0,6251**	0,2580	0,1534	0,6614***	0,2447	0,1623
Escol_professor_ens_superior	-0,0381	0,1752	-0,0095	-0,0889	0,1727	-0,0221
Experiência_professor	-0,1530	0,1312	-0,0381	-0,1237	0,1329	-0,0308
Intercepto	-1,9722***	0,5157	—	-3,6078***	0,5982	—
<b>Número de observações</b>	1.193					

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Notas: Valores expandidos pelo peso amostral. \*Estatisticamente significativo a 10%; \*\*Estatisticamente significativo a 5%; \*\*\*Estatisticamente significativo a 1%. A variável Banheiros\_escola\_bom não foi utilizada no cálculo do escore de propensão da prova de português para que a propriedade de suporte comum pudesse ser satisfeita. No caso de matemática, foi necessário omitir a variável Escol\_diretor\_ens\_superior.

Os coeficientes estimados na Tabela 2 indicam que, de modo geral, o fato do aluno ser homem reduz a probabilidade em quase 14 pontos percentuais (p.p.) dele obter um desempenho acima da média na prova de português. Já em matemática, o resultado é mais favorável aos estudantes do sexo masculino, indicando que estes possuem uma probabilidade maior de obter um desempenho acima da média. Em ambas as disciplinas, a variável que apresentou efeito marginal mais elevado foi aquela referente ao incentivo dos pais ao estudo: cerca de 20 p.p. em português e 41 p.p. em matemática. Outros resultados expressivos foram captados pela variável

que indica se o aluno trabalha ou não fora de casa, pela variável que diz respeito à ocorrência de reprovação por parte do aluno, pela variável de experiência do diretor e pela variável que indica se o aluno iniciou os estudos na creche ou pré-escola<sup>17</sup>.

Com relação às variáveis de esforço, é possível observar que o hábito dos alunos de fazer a tarefa de casa apresentou significância estatística apenas para português, com efeito marginal positivo (8,39 p.p.). Além disso, ao contrário do esperado, a variável de hábito de leitura apresentou efeito marginal negativo. Todavia, como será visto mais adiante, essa mesma variável mostrou-se diretamente relacionada com a proficiência em português para os alunos do 9º ano.

Os resultados também mostraram que as características da escola não parecem ter relação direta significativa com o desempenho dos estudantes. Entretanto, tais efeitos já eram esperados, pois os mesmos já foram observados em várias referências da literatura<sup>18</sup>. Apenas a existência de uma boa biblioteca, em relação à proficiência em matemática, apresentou impacto positivo e estatisticamente significativo.

A escolaridade do professor, bem como sua experiência de ensino, não apresentou significância nem para português e nem para matemática. Uma possível explicação para os efeitos observados com essas variáveis foi dada por Hanushek (2007), que afirma que não necessariamente serão os professores mais experientes ou mais capacitados que afetarão o desempenho dos alunos, mas sim aqueles com melhor didática e maior esforço ao preparar suas aulas.

Adiante, na Tabela 3, encontram-se os coeficientes estimados para os alunos do 9º ano. De forma similar para com os estudantes do 5º ano, o modelo indica que os estudantes do sexo feminino se sobressaem em português, enquanto que os estudantes do sexo masculino apresentam melhor desempenho em matemática. Diferentemente dos resultados anteriores, a presença no domicílio de um computador com acesso à internet, e também a experiência dos professores, encontram-se positivamente relacionadas com o desempenho dos discentes, em ambas as proficiências. Também vale ressaltar que a variável de incentivo dos pais ao estudo, que mostrou-se importante para um bom desempenho dos estudantes do 5º ano, não apresentou significância estatística para os modelos referentes aos alunos do 9º ano.

---

<sup>17</sup> A importância da creche e da pré-escola na proficiência escolar foi analisada por Curi e Menezes-Filho (2006).

<sup>18</sup> Ver Hanushek (2007) para uma melhor compreensão do impacto da escola e dos professores sobre o desempenho dos estudantes.



**Tabela 3 – Estimação do escore de propensão (9º Ano)**

Variáveis Explicativas	Português			Matemática		
	Coefficiente	Erro-padrão	dy/dx	Coefficiente	Erro-padrão	dy/dx
Sexo_aluno_masculino	-0,2937***	0,0760	-0,0722	0,5888***	0,0774	0,1454
Raça_aluno_branco	0,0614	0,0863	0,0150	-0,0223	0,0863	-0,0056
Computador_com_internet	0,1520**	0,0759	0,0373	0,1555**	0,0757	0,0387
Escol_mãe_ens_superior	0,0934	0,1283	0,0228	0,1478	0,1286	0,0366
Incentivo_aluno	0,3518	0,3208	0,0875	0,0318	0,3112	0,0079
Ler_livros_sempre	0,2450***	0,0849	0,0597	0,1095	0,0840	0,0272
Trabalho_aluno	-0,4867***	0,1195	-0,1209	-0,3022**	0,1192	-0,0754
Creche_ou_pré-escola	0,3541***	0,1043	0,0879	0,2709**	0,1052	0,0676
Reprovação_aluno	0,6421***	0,0757	0,1573	0,7576***	0,0764	0,1870
Faz_tarefa_sempre	0,0661	0,0767	0,0163	0,4847***	0,0746	0,1202
Prof_corrige_tarefa_sempre	0,1550*	0,0868	0,0382	—	—	—
Salas_escola_bom	0,1072	0,1010	0,0264	0,0146	0,1017	0,0036
Banheiros_escola_bom	0,0686	0,0930	0,0168	0,1070	0,0926	0,0266
Biblioteca_escola_bom	0,0112	0,1023	0,0027	0,0508	0,1018	0,0127
Laboratório_escola_bom	0,1523*	0,0807	0,0374	-0,1639**	0,0805	-0,0409
Escol_diretor_ens_superior	-0,1427	0,2622	-0,0347	-0,2925	0,2683	-0,0718
Experiência_diretor	-0,1719	0,3288	-0,0417	-0,1381	0,3139	-0,0342
Escol_professor_ens_superior	0,1889	0,5045	0,0468	-0,3793	0,5286	-0,0924
Experiência_professor	0,1913**	0,0778	0,0471	0,2083***	0,0779	0,0519
Intercepto	-1,0685	0,7601	—	-0,5289	0,7657	—
<b>Número de observações</b>	<b>2.486</b>					

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Notas: Valores expandidos pelo peso amostral. \*Estatisticamente significante a 10%; \*\*Estatisticamente significante a 5%; \*\*\*Estatisticamente significante a 1%. A variável Prof\_corrige\_tarefa\_sempre não foi utilizada no cálculo do escore de propensão da prova de matemática para que a propriedade de suporte comum pudesse ser satisfeita.

Para a disciplina de português, as variáveis que apresentaram maiores efeitos marginais foram as relacionadas à questão da reprovação e ao fato do estudante trabalhar ou não. Com relação à primeira variável, os resultados apontaram que caso o aluno nunca tenha reprovado, sua probabilidade de obter um desempenho acima da média aumenta em quase 16 p.p. Já em relação à segunda variável, o efeito é de uma redução de pouco mais de 12 p.p. para aqueles que trabalham. Em matemática, a variável de reprovação também apresentou o maior efeito marginal (18,70 p.p.), sendo seguida pelas variáveis de sexo do aluno e de resposta da tarefa de casa.

Também é importante destacar que, novamente, as variáveis relacionadas às características das escolas não se mostraram relevantes para o desempenho dos alunos. Apenas a existência de um bom laboratório apresentou impacto significativo, sendo este impacto

positivo em português e negativo em matemática. Outro resultado interessante foi encontrado pela variável de experiência do professor, que não mostrou ser importante para os estudantes do 5º ano. No caso dos discentes do 9º ano, esta variável foi significativa para ambas as disciplinas e apresentou efeito marginal positivo.

Por fim, tendo em vista os resultados dos modelos estimados, conclui-se que as seguintes variáveis são as principais determinantes para que os estudantes apresentem desempenho acima da média: sexo do aluno, presença de computador com internet no domicílio, incentivo dos pais ao estudo, hábito de leitura, se o aluno trabalha, o momento em que o aluno iniciou seus estudos, se o aluno reprovou alguma vez, se o aluno costuma fazer a tarefa de casa, a experiência do diretor da escola e a experiência do professor. Tais resultados estão em consonância com aqueles encontrados em Barros et al (2001), Felício e Fernandes (2005), Curi e Menezes-Filho (2006); Menezes-Filho (2007), Machado et al. (2008), Vernier e Bagolin (2013) e Almeida (2014), bem como em outros trabalhos da literatura. No cálculo da decomposição de Oaxaca-Blinder serão essas as variáveis explicativas consideradas<sup>19</sup>.

#### **2.4.2 ESTIMAÇÃO DA DECOMPOSIÇÃO DE OAXACA-BLINDER**

Para decompor o diferencial de desempenho escolar em fatores observáveis e não observáveis, empregou-se o procedimento de Oaxaca-Blinder (1973) construindo-se grupos amostrais a partir das seguintes informações: se o aluno iniciou os estudos na creche ou pré-escola, se os pais do aluno o incentivam a estudar, se o aluno faz a tarefa de casa sempre ou quase sempre, se o aluno nunca foi reprovado e se o aluno apresentou desempenho acima da média. A escolha por esses grupos se justifica pela importância que tais variáveis têm apresentado em diversos estudos que buscam identificar os determinantes do desempenho escolar (BARROS et al., 2001; SOUZA, 2005; CURI e MENEZES-FILHO 2006; MENEZES-FILHO, 2007; ARAÚJO e ALMEIDA, 2013).

Na Tabela 4 é possível observar os resultados da decomposição de Oaxaca-Blinder sem pareamento. Primeiramente, nota-se que o grupo amostral construído pela variável acerca da creche ou pré-escola apresentou significância estatística em todos os resultados. De acordo com esse modelo, os estudantes do 5º ano que iniciaram seus estudos na creche ou pré-escola apresentam, em média, desempenho superior em aproximadamente 10,90 pontos. É importante observar que a maior parte desse diferencial é atribuído aos fatores não explicados, isto é, não

---

<sup>19</sup> Em razão disso, o número de observações nas estimações da decomposição de Oaxaca-Blinder com e sem PSM diferem daqueles apresentados na seção de fonte de tratamento dos dados.

observáveis diretamente<sup>20</sup>. Já para os estudantes do 9º ano, o diferencial é um pouco maior: 12,30 pontos em português e 10,28 em matemática, onde novamente a maior parcela dos diferenciais é decorrente de características não observáveis.

**Tabela 4 – Resultados da decomposição de Oaxaca-Blinder**

<b>O aluno iniciou os estudos na creche ou pré-escola<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	10,8994***	9,5090***	12,2993***	10,2805***
Diferença explicada	3,4732***	2,7788***	3,7059***	3,8091***
Diferença não explicada	7,4262***	6,7303***	8,5934***	6,4715***
<b>Os pais do aluno o incentivam a estudar<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	21,4681***	30,0724***	18,0627***	9,5160
Diferença explicada	3,2929**	2,0563*	6,4453***	5,4310***
Diferença não explicada	18,1751***	28,0161***	11,6174*	4,0850
<b>O aluno faz a tarefa de casa sempre ou quase sempre<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	11,3100***	7,1539***	8,8566***	16,2553***
Diferença explicada	3,4911***	1,3900**	5,2141***	2,8862***
Diferença não explicada	7,8189***	5,7637***	3,6425**	13,3691***
<b>O aluno nunca foi reprovado<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	19,7227***	16,0571***	21,7006***	22,2098***
Diferença explicada	3,7622***	1,5008***	3,6748***	1,9346***
Diferença não explicada	15,9606***	14,556***	18,0258***	20,2752***
<b>O aluno apresentou desempenho acima da média<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	65,2998***	64,0573***	72,9445***	71,6473***
Diferença explicada	2,3558***	2,0829***	2,5472***	3,2824***
Diferença não explicada	62,9440***	61,9744***	70,3973***	68,3648***
<b>Número de observações</b>	<b>2.558</b>	<b>2.571</b>	<b>3.423</b>	<b>3.428</b>

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Notas: Valores expandidos pelo peso amostral. \*Estatisticamente significativo a 10%; \*\*Estatisticamente significativo a 5%; \*\*\*Estatisticamente significativo a 1%. (\*) Valores positivos para as diferenças indicam vantagem para os estudantes que iniciaram os estudos na creche ou pré-escola, para os estudantes cujos pais o incentivam a estudar, para os alunos que fazem a tarefa de casa sempre ou quase sempre, para aqueles que nunca reprovaram e para os discentes que apresentaram desempenho acima da média. Valores negativos indicam vantagens para os casos contrários.

<sup>20</sup> Lembrando que os fatores observáveis considerados aqui são os seguintes: sexo do aluno, presença de computador com internet no domicílio, incentivo dos pais ao estudo, hábito de leitura, se o aluno trabalha, se o aluno iniciou os estudos na creche ou pré-escola, se o aluno reprovou alguma vez, se o aluno costuma fazer a tarefa de casa, a experiência do diretor e a experiência do professor.

Na segunda estimação – utilizando a variável de incentivo dos pais ao estudo como forma de segregar os estudantes – os resultados sugerem que tal incentivo apresenta impactos positivos no desempenho escolar. Na proficiência em português, tanto no 5º quanto no 9º ano, o diferencial de desempenho calculado ficou em torno de 20 pontos. Em matemática, apenas o resultado do 5º ano mostrou-se estatisticamente significativo, indicando um diferencial de 30,07 pontos. Mais uma vez, é notório o importante papel desempenhado pelos fatores não explicados.

Os resultados do terceiro modelo indicam que os discentes que costumam fazer a tarefa de casa tendem a apresentar melhor desempenho do que aqueles que não fazem-na com tanta frequência, ou nunca a fazem. Para os alunos do 5º ano, na prova de português, a diferença total encontrada foi de 11,31 pontos, em que quase 8 pontos foram atribuídos aos fatores não explicados. Um resultado que merece destaque é o diferencial observado no desempenho da prova de matemática dos alunos do 9º ano, cerca de 16,25 pontos. Novamente, os resultados apontam para uma maior importância relativa dos componentes não observáveis.

Mais adiante, diferenciando os alunos entre os que nunca reprovaram e os que já reprovaram ao menos uma vez, nota-se que os alunos do primeiro grupo apresentam um melhor desempenho tanto em português quanto em matemática, para ambos os anos. Da mesma forma como foi observado nos demais modelos, a parcela do diferencial total decorrente dos fatores não observáveis é majoritariamente dominante.

Por fim, o último modelo estimado buscou mensurar o diferencial de desempenho entre os estudantes acima e abaixo da média, verificando que esse diferencial é decorrente não apenas dos fatores observáveis, mas também dos fatores não observáveis. Como pode ser visto na Tabela 4, todos os resultados desse modelo mostraram-se estatisticamente significativos. Para os estudantes do 5º ano, os diferenciais de desempenho nas provas de português e matemática foram de, respectivamente, 65,30 e 64,06 pontos. Já para os estudantes do 9º ano, os resultados foram de 72,94 e 71,65 pontos. Em ambos os anos, praticamente toda a parcela dos diferenciais pode ser atribuída aos fatores não explicados.

Os resultados obtidos com os cinco modelos acima descritos corroboraram com o argumento apresentado neste artigo de que a diferença de desempenho existente entre boa parte dos estudantes é decorrente do efeito contingência, ou seja, de características particulares, inatas, que geralmente não podem ser observadas diretamente. Isso fica evidente ao se observar que mesmo controlando e considerando variáveis determinantes do desempenho escolar, o que inclui também as *proxies* de esforço e incentivo, a maior parte dos diferenciais ainda recai sobre os fatores não observáveis.

### 2.4.3 ESTIMAÇÃO DA DECOMPOSIÇÃO DE OAXACA-BLINDER COM PSM

Conforme ressaltado anteriormente, é possível tornar a amostra mais homogênea empregando a técnica de Pareamento por Escore de Propensão<sup>21</sup>. Dessa forma, as comparações realizadas são feitas entre estudantes com características observáveis similares, tornando os resultados encontrados mais robustos. Na Tabela 5 encontram-se os resultados da decomposição de Oaxaca-Blinder com PSM, considerando as mesmas variáveis explicativas e de segregação adotadas nos modelos da Tabela 4.

**Tabela 5 – Resultados da decomposição de Oaxaca-Blinder com PSM**

<b>O aluno iniciou os estudos na creche ou pré-escola<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	12,3777***	11,6227***	11,6725***	11,6768***
Diferença explicada	3,9302***	4,1473***	4,0470***	4,0812***
Diferença não explicada	8,4474***	7,4754**	7,6255***	7,5956***
<b>Os pais do aluno o incentivam a estudar<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	22,6943***	35,8144***	18,1918**	11,8829
Diferença explicada	2,8421	-2,3997	3,2793	5,7323**
Diferença não explicada	19,8522 ***	38,2142***	14,9124*	6,1506
<b>O aluno faz a tarefa de casa sempre ou quase sempre<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	9,5694***	6,6885**	7,4174***	17,1510***
Diferença explicada	3,0187***	1,4715	5,0907***	2,9957***
Diferença não explicada	6,5507**	5,2170*	2,3267	14,1554***
<b>O aluno nunca foi reprovado<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	20,0640***	18,1083***	21,0742***	23,5566***
Diferença explicada	3,6720***	1,7550**	3,4456***	1,9444***
Diferença não explicada	16,3920***	16,3533***	17,6286***	21,6122***
<b>O aluno apresentou desempenho acima da média<sup>(*)</sup></b>				
	<b>5º Ano</b>		<b>9º Ano</b>	
	<b>Português</b>	<b>Matemática</b>	<b>Português</b>	<b>Matemática</b>
Diferença total	67,5285***	66,3283***	71,3730***	71,5603***
Diferença explicada	1,9076***	2,3510***	2,6177***	3,7579***
Diferença não explicada	65,6209***	63,9773***	68,7553***	67,8025***

<sup>21</sup> Vale salientar que não é necessário calcular os escores de propensão para cada um dos modelos estimados na Tabela 5. Como o objetivo do PSM é parear a amostra, criando uma região de suporte comum e homogeneizando-a, é possível usar os mesmos escores calculados anteriormente em todas as estimações.

Número de observações	1.194	1.209	2.503	2.504
-----------------------	-------	-------	-------	-------

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Notas: Valores expandidos pelo peso amostral. \*Estatisticamente significativo a 10%; \*\*Estatisticamente significativo a 5%; \*\*\*Estatisticamente significativo a 1%. (\*)Valores positivos para as diferenças indicam vantagem para os estudantes que iniciaram os estudos na creche ou pré-escola, para os estudantes cujos pais os incentivam a estudar, para os alunos que fazem a tarefa de casa sempre ou quase sempre, para aqueles que nunca reprovaram e para os discentes que apresentaram desempenho acima da média. Valores negativos indicam vantagens para os casos contrários.

De modo geral, os resultados da decomposição de Oaxaca-Blinder com e sem PSM foram muito similares, com pequenas diferenças na magnitude dos coeficientes. Comparando o primeiro modelo de ambas as decomposições, nota-se que com PSM o diferencial total nas notas de português e matemática dos alunos do 5º ano é mais elevado. Com relação aos alunos do 9º ano, o diferencial foi maior apenas em matemática. Dessa forma, os resultados indicam que aqueles alunos que iniciaram seus estudos na creche ou pré-escola apresentam, em média, desempenho superior de aproximadamente 12 pontos, sendo que grande parcela desse diferencial é ocasionado por fatores não explicados.

No segundo modelo, em que a divisão dos grupos foi feita por meio da variável de incentivo dos pais aos estudos, observa-se que a parcela do diferencial total atribuída aos fatores não observáveis foi maior no modelo com PSM. Também é interessante observar que, em ambos os casos, o coeficiente da diferença total da disciplina de matemática para os alunos do 9º ano não apresentou significância estatística, o que sugere que o desempenho nessa matéria não é diretamente afetado pelo incentivo dos pais.

Com relação ao terceiro modelo, cuja variável de controle usada para separar os grupos foi referente ao esforço dos alunos em fazer a tarefa de casa, tem-se que a diferença total foi menor para os casos com PSM, à exceção de matemática para estudantes do 9º ano. Foi percebido, conforme já esperado, a existência de um diferencial favorável aos estudantes que costumam fazer suas tarefas, em que a maior parte desse diferencial, em quase todas as análises, decorre principalmente de fatores não explicados.

O quarto modelo separou a amostra com base na ocorrência ou não de reprovação por parte dos alunos. Todos os coeficientes desse modelo apresentaram significância estatística, e os maiores diferenciais totais encontrados deram-se para os alunos do 9º ano. Em média, os resultados dos alunos que nunca foram reprovados é mais de 20 pontos superior em relação aos que já foram reprovados. Novamente, a maior parte do diferencial total recai sobre as características não observáveis.

Por último, tem-se o modelo que busca decompor a diferença de desempenho entre os estudantes que obtiveram desempenho acima e abaixo da média, considerando apenas os

estudantes dentro da região de suporte comum. Os resultados foram similares ao modelo sem PSM, indicando que grande parte do diferencial existente é devido aos fatores não explicados.

Como pôde ser visto nesta seção de resultados, a decomposição de Oaxaca-Blinder com PSM também apresenta evidências favoráveis ao efeito contingência descrito por Rawls (1971), como forma de explicar o diferencial de desempenho entre os discentes. Isso fica claro pois, mesmo controlando as comparações para que essas fossem realizadas entre indivíduos similares, e considerando o efeito da escola, do domicílio, dos professores, do diretor e incluindo variáveis de esforço e incentivo, ainda existe um diferencial não explicado que é oriundo das particularidades e capacidades inatas de cada estudante.

## 2.5 CONSIDERAÇÕES FINAIS

São vários os trabalhos na literatura que buscam identificar os determinantes do desempenho escolar e a maior parte dessas pesquisas identifica nas características individuais dos alunos, bem como no *background* familiar, os principais fatores que explicam os resultados observados. Mais recentemente, as questões de justiça e equidade vêm sendo debatidas nesse contexto, e o papel da desigualdade de oportunidade tem recebido cada vez mais espaço. Todavia, apesar de todo o avanço já realizado, ainda não há consenso sobre como seria possível oferecer educação de qualidade de modo mais abrangente e justo.

Rawls (1971) apresentou uma teoria considerando que alguns indivíduos são, por natureza, diferentes em suas dotações, isto é, diferentes em suas habilidades e talentos natos. Desse modo, existiria uma desigualdade de contingência que explicaria o porquê de indivíduos com acesso aos mesmos recursos e benefícios apresentarem resultados tão diferentes. Tendo isso em vista, o presente estudo teve por objetivo verificar a ocorrência de desigualdade de contingência entre os alunos do 5º e 9º ano das escolas públicas de João Pessoa, a partir do desempenho em português e matemática na Prova Brasil de 2011.

Como estratégia empírica, adotou-se a decomposição de Oaxaca-Blinder combinada com o método de Pareamento por Escore de Propensão. Desse modo, foi possível garantir que as comparações fossem realizadas apenas entre estudantes com características observáveis semelhantes.

A partir dos resultados aferidos pela decomposição de Oaxaca-Blinder foram encontradas evidências favoráveis à existência de desigualdade de contingência entre os alunos do 5º e 9º ano, tanto em português quanto em matemática, uma vez que se observaram diferenças de mais de 66 pontos, na média, atribuídas majoritariamente às características não

observáveis. Entretanto, vale frisar que mesmo considerando as *proxies* de esforço e incentivo ainda há fatores que não puderam ser considerados na análise devido à ausência de informações. Como exemplos desses fatores é possível citar a herdabilidade genética e a questão do ambiente compartilhado, que podem influenciar o desempenho dos estudantes e estariam sendo captadas nos fatores não observados<sup>22</sup>.

Finalmente, este trabalho se propôs a contribuir com a literatura da economia da educação atribuindo um novo conceito às características não observáveis. Os resultados encontrados indicam que o diferencial de desempenho entre os estudantes pesquisados pode ser, pelo menos em parte, atribuído à desigualdade de contingência, ou seja, às diferenças nas habilidades inatas. Sendo assim, há razões para acreditar que mesmo em situações onde não haja desigualdade de oportunidade, ainda assim será observado diferenças no desempenho desses discentes.

---

<sup>22</sup> Ver, por exemplo, os trabalhos de Chipuer, Rovine e Plomin (1990), Herskind et al. (1996), Ronald, Spinath e Plomin (2002), Plomin e Spinath (2004).



## REFERÊNCIAS

- ALMEIDA, A. T. C. Determinantes dos piores e melhores resultados educacionais dos alunos da rede pública de Ensino Fundamental no Brasil. **Planejamento e Políticas Públicas**, v. 42, p. 147-188, 2014.
- ARAÚJO, E. S. P.; ALMEIDA, A. T. C. Avaliação dos resultados educacionais dos alunos das escolas municipais de João Pessoa-PB. **MPGOA**, v. 2, p. 46-63, 2013.
- BARROS, R. P.; MENDONÇA, R.; SANTOS, D. D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, v. 31, n. 1, p. 1-42, abr. 2001.
- BECKER, S. O.; ICHINO, A. Estimation of average treatment effects based on propensity scores. **The Stata Journal**, v. 2, n. 4, p. 358–377, 2002.
- BLINDER, A. S. Wage discrimination: reduced form and structural estimates. **Journal of Human Resources**, v.8, n. 4, p. 436–455, 1973.
- BRASIL. Ministério da Educação e Cultura / Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Microdados Prova Brasil 2011**: manual do usuário. Brasília-DF: INEP, 2012.
- CHIPUER, H. M.; ROVINE, M. J.; PLOMIN, R. LISREL modeling: Genetic and environmental influences on IQ revisited. **Intelligence**, v. 14, n. 1, p. 11-29, 1990.
- CURI, A. Z.; MENEZES FILHO, N. A. Os efeitos da pré-escola sobre os salários, a escolaridade e a proficiência escolar. In: XXXIV Encontro Nacional de Economia, 2006, Salvador. Anais do XXXIV Encontro Nacional de Economia. São Paulo: ANPEC, 2006. v. 1.
- FELÍCIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo. In: ENCONTRO NACIONAL DE ECONOMIA DA ANPEC, 2005, Natal, Rio Grande do Norte. Anais do XXXIII Encontro Nacional de Economia. Natal: ANPEC, 2005.
- FIGUEIREDO, E.; SILVA, C. R. F. E. Desigualdade de Oportunidades no Brasil: Uma Decomposição Quantílica Contrafactual. **Pesquisa e Planejamento Econômico** (Rio de Janeiro), v. 42, p. 29-48, 2012.
- GONDIM, E. John Rawls: a educação política. **Revista Thaumazein (Santa Maria)**, v. 5, p. 1-15, 2010.
- GONDIM, E.; RODRIGUES, O. M. John Rawls: educação, cidadania e equilíbrio reflexivo. **Revista Filosofia Capital**, v. 6, p. 25-33, 2011.
- HANUSHEK, E. A. **Education Production Functions**. Stanford: Stanford University, January 2007.

HANUSHEK, E. A. **Do Teachers Make a Difference?** In: U.S. Office of Education. The Production of Education, Teacher Quality, and Efficiency. Washington, D.C.: Government Printing Office, 1970, pp. 79-99.

HERSKIND, A. M.; HOLM, M. M. N. V.; SØRENSEN, T. I. A.; VAUPEL, B. H. J. W. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. **Human genetics**, v. 97, n. 3, p. 319-323, 1996.

HESHMATI, A. **Inequalities and their measurement**. Unpublished manuscript, 2004.

JANN, B. The Blinder–Oaxaca decomposition for linear regression models. **The Stata Journal**, v. 8, n. 4, p. 453–479, 2008.

KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. **Handbook on impact evaluation: Quantitative Methods and Practices**. Washington, DC: The World Bank, 2010.

KRANICH, L. Equitable opportunities: an axiomatic approach. **Journal of Economic Theory**, v. 71, p. 132-147, 1996.

MACHADO, A. F.; MORO, S.; RODRIGUES, L.; RIOS, J. Qualidade do Ensino em Matemática: determinantes do desempenho de alunos em escolas públicas estaduais mineiras. **Revista ANPEC**, v. 9, p. 23-46, 2008.

MENEZES FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012.

MENEZES FILHO, N. Os Determinantes do Desempenho Escolar no Brasil. Instituto Futuro Brasil, São Paulo, Ibmecc – SP, 2007.

ÑOPO, H. Matching as a tool to decompose wage gaps. **The Review of Economics and Statistics**, v. 90, n. 2, p. 290-299, 2008.

OAXACA, R. Male–female wage differential in urban labor markets. **International Economic Review**, v. 14, n. 3, p. 693–709, 1973.

PATTANAIK, P. K.; XU, Y. On ranking opportunity sets in terms of freedom of choice. **Recherches Économiques de Louvain**, v. 56, n. 3/4, 1990.

PLOMIN, R.; SPINATH, F. M. Intelligence: genetics, genes, and genomics. **Journal of Personality and Social Psychology**, v. 86, n. 1, p. 112, 2004.

RAWLS, J. **A theory of justice**. United States of America: Harvard University Press, 1971.

RONALD, A.; SPINATH, F. M.; PLOMIN, R. The aetiology of high cognitive ability in early childhood. **High Ability Studies**, v. 13, n. 2, p. 103-114, 2002.

ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, n. 1, p. 41–55, 1983.

SEN, A. **Commodities and Capabilities**. Amsterdam: North-Holland, 1985.

SOUZA, A. M. Determinantes da aprendizagem em escolas municipais. **Ensaio. Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 13, n.48, p. 271-292, 2005.

THOMSON, W. Fair allocation rules. In: ARROW, K.; SEN, A.; SUZUMURA, K. (Ed.). **Handbooks in economics: social choice and welfare**. Amsterdam: Elsevier, 2011.

VERNIER, L. D. S.; BAGOLIN, I. P. Determinantes do desempenho escolar no estado do Rio Grande do Sul: uma análise com regressões quantílicas. In: XVI ENCONTRO DE ECONOMIA DA REGIÃO SUL, 2013, Curitiba. XVI Encontro de Economia da Região Sul, 2013. v. 1. p. 1-20.

## APÊNDICE

Tabela 6 - Estatísticas descritivas da base de dados completa para os alunos do 5º e 9º ano do município de João Pessoa

	Variáveis	5º ano					9º ano				
		Obs.	Média	Desvio-padrão	Mínimo	Máximo	Obs.	Média	Desvio-padrão	Mínimo	Máximo
<b>Características do Aluno e do Domicílio</b>	Sexo_aluno_masculino	5.878	0,5092	0,5000	0	1	4.668	0,4680	0,4990	0	1
	Raça_aluno_branco	4.954	0,2778	0,4480	0	1	4.325	0,2442	0,4297	0	1
	Trabalho_aluno	5.780	0,1387	0,3456	0	1	4.651	0,1231	0,3285	0	1
	Incentivo_aluno	5.583	0,9652	0,1833	0	1	4.597	0,9854	0,1199	0	1
	Ler_livros_sempre	5.228	0,4180	0,4933	0	1	4.279	0,2703	0,4441	0	1
	Faz_tarefa_português_sempre	5.906	0,7302	0,4439	0	1	4.664	0,5429	0,4982	0	1
	Faz_tarefa_matemática_sempre	5.926	0,7344	0,4417	0	1	4.667	0,4981	0,5000	0	1
	Creche_ou_pré-escola	5.807	0,7524	0,4317	0	1	4.628	0,8466	0,3604	0	1
	Reprovação_aluno	5.911	0,5877	0,4923	0	1	4.684	0,5481	0,4977	0	1
	Computador_com_internet	6.022	0,3610	0,4803	0	1	4.694	0,4861	0,4998	0	1
	Escol_mãe_ens_superior	3.630	0,1551	0,3621	0	1	3.905	0,0885	0,2841	0	1
	Prof_corrige_tarefa_port_sempre	5.886	0,7686	0,4217	0	1	4.686	0,7425	0,4373	0	1
	Prof_corrige_tarefa_mat_sempre	5.956	0,7904	0,4070	0	1	4.679	0,7516	0,4321	0	1
	Proficiência_português	6.172	184,4041	40,4947	83,4419	327,4591	4.753	238,0066	45,1266	106,0994	360,7970
	Proficiência_matemática	6.172	199,8191	39,4889	99,6468	337,1514	4.753	239,7518	44,4367	112,6782	396,0480
<b>Características da Escola</b>	Salas_escola_bom	5.991	0,7767	0,4165	0	1	4.735	0,7411	0,4381	0	1
	Banheiros_escola_bom	6.107	0,5709	0,4950	0	1	4.730	0,4509	0,4976	0	1
	Biblioteca_escola_bom	6.119	0,7199	0,4491	0	1	4.710	0,7870	0,4095	0	1
	Laboratório_escola_bom	5.957	0,4645	0,4988	0	1	4.548	0,4186	0,4934	0	1
<b>Características do Professor</b>	Escol_professor_ens_superior	3.716	0,8598	0,3472	0	1	4.155	0,9943	0,0755	0	1
	Experiência_professor	3.759	0,7235	0,4473	0	1	4.113	0,6147	0,4867	0	1
<b>Características do Diretor</b>	Escol_diretor_ens_superior	6.027	0,9507	0,2164	0	1	4.753	0,9791	0,1430	0	1
	Experiência_diretor	6.027	0,9625	0,1901	0	1	4.753	0,9812	0,1359	0	1

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2011.

Nota: Valores expandidos pelo peso amostral.

### **3. AVALIAÇÃO DE IMPACTO DO PROGRAMA MAIS EDUCAÇÃO NAS ESCOLAS DA REDE PÚBLICA DE ENSINO FUNDAMENTAL: UMA ANÁLISE PARA O PERÍODO 2007-2011**

#### **3.1 INTRODUÇÃO**

No Brasil, ao longo dos anos 1990, a educação passou a ganhar mais evidência na agenda governamental, sobretudo a partir da sanção da Lei de Diretrizes e Bases da Educação Nacional, em 1996. Essa lei veio para reafirmar o direito à educação, garantido pela Constituição Federal, estabelecendo os princípios e deveres do Estado em relação à educação escolar pública. Ainda nesse período, houve também a criação do Fundo de Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (FUNDEF), do Programa Dinheiro Direto na Escola (PDDE), do Programa Bolsa-Escola e do Programa de Erradicação do Trabalho Infantil (PETI), dentre outros projetos sociais.

Nos anos 2000, com a economia mais estabilizada<sup>23</sup>, houve uma significativa expansão dos gastos governamentais em praticamente todas as áreas sociais. Na educação, além de ter sido observado a expansão dos recursos destinados a uma série de programas já existentes, vale destacar as ações articuladas no Plano de Desenvolvimento da Educação (PDE), lançado em 2007, a substituição do FUNDEF pelo Fundo de Desenvolvimento da Educação Básica e Valorização dos Profissionais da Educação (FUNDEB) e a criação dos programas Brasil Alfabetizado e Mais Educação.

No esteio desse processo, o Programa Mais Educação foi instituído em 2007, e, a partir de 2008, tem ampliado de forma considerável sua cobertura. O programa adota a perspectiva da educação em tempo integral e suas ações acontecem no contra turno escolar. Sendo assim, pode-se dizer que ele atua mantendo crianças e jovens mais tempo na escola, tornando-as menos suscetíveis às situações de violência e fornecendo uma formação mais completa e cidadã<sup>24</sup>.

Vários estudos têm sido realizados para analisar pedagogicamente o funcionamento do Mais Educação e constatar se seus objetivos vêm sendo alcançados<sup>25</sup>. Entretanto, são poucos os trabalhos que empregaram métodos de avaliação de impacto para quantificar seus efeitos, com destaque para Pereira (2011), que obteve estimativas para Minas Gerais e para o Brasil, Mendes (2011), que avaliou o impacto sobre as notas e a taxa de aprovação dos alunos do

---

<sup>23</sup> Principalmente em função do Plano Real, implementando em 1994.

<sup>24</sup> Por formação cidadã entende-se a construção de cidadãos mais conscientes de sua cultura, da diversidade étnica de seu país, das preocupações com as questões ambientais e sociais, etc.

<sup>25</sup> Vide Lopes e Araújo (2012), Pires et al. (2012), Garcia (2013) e Nóbrega e Silva (2013).

Ensino Fundamental da rede pública brasileira, e Xerxenevsky (2012), que analisou os efeitos do programa sobre o desempenho de alunos do Rio Grande do Sul.

Diante disso, o objetivo do presente estudo é avaliar o impacto do Programa Mais Educação nas escolas de Ensino Fundamental da rede pública do estado da Paraíba, dos estados da região Nordeste e do Brasil, visando comparar os resultados encontrados. Apesar de já existirem outros trabalhos que avaliaram o referido programa para outras localidades, esta pesquisa inova pela dimensão temporal<sup>26</sup> e também por contemplar cinco variáveis de impacto: o Índice de Desenvolvimento da Educação Básica (IDEB) das escolas<sup>27</sup>, a nota média das escolas em português e em matemática, a taxa de abandono e o índice de violência das escolas. A escolha por esses indicadores justifica-se pelo próprio desenho do programa, haja visto a extensa gama de atividades contempladas.

Para alcançar os objetivos propostos são utilizados os microdados da Prova Brasil de 2007, 2009 e 2011. Como as atividades do programa tiveram início apenas em 2008, é possível construir um painel contendo informações acerca das escolas em um período anterior e em dois períodos posteriores ao início do Mais Educação. Vale ressaltar que escolha temporal justifica-se pelo tempo necessário para a maturação do programa, de modo que o mesmo possa atingir os objetivos propostos.

A estrutura de painel aliada ao conjunto de características da escola, dos alunos, dos professores e dos diretores, permite combinar o método de Diferenças em Diferenças com o método de Pareamento por Escore de Propensão. A adoção dessa estratégia empírica fornece estimativas robustas do efeito do programa sobre os indicadores de impacto, pois controla tanto pelo efeito temporal quanto pelas características observáveis.

Além dessa introdução, essa pesquisa contém mais cinco seções. A próxima seção discute o funcionamento do programa Mais Educação. Em seguida, é apresentada a fonte e o tratamento dos dados utilizados, sendo acompanhada pela descrição da estratégia empírica adotada. A quarta seção apresenta a análise dos resultados e, por fim, a quinta e última seção contempla as considerações finais.

---

<sup>26</sup> Os trabalhos de Pereira (2011), Mendes (2011) e Xerxenevsky (2012) estimaram os impactos do Mais Educação utilizando os microdados da Prova Brasil de 2007 e 2009, enquanto que nessa pesquisa serão utilizados os microdados de 2007, 2009 e 2011

<sup>27</sup> O IDEB é resultado da combinação de dois indicadores educacionais, a saber: as pontuações em testes padronizados (como a Prova Brasil, por exemplo) e informações sobre rendimento escolar (taxa de aprovação, repetência e evasão).

### 3.2 PROGRAMA MAIS EDUCAÇÃO

O Programa Mais Educação foi instituído pela Portaria Interministerial Nº 17/2007 e regulamentado pelo Decreto Presidencial 7.083/2010, sendo parte constitutiva do Plano de Desenvolvimento da Educação. Implementado efetivamente em 2008, o referido programa atende prioritariamente às escolas com baixo IDEB situadas em capitais e regiões metropolitanas, buscando oferecer a crianças e jovens a ampliação da jornada escolar para no mínimo sete horas diária e a organização curricular na perspectiva da educação integral (MINISTÉRIO DA EDUCAÇÃO, 2009).

De acordo com o Artigo 2º da Portaria Normativa Interministerial nº 17, de 24 de abril de 2007, o Programa Mais Educação tem como finalidades:

- i) Apoiar a ampliação do tempo e do espaço educativo e a extensão do ambiente escolar nas redes públicas de educação básica de estados, Distrito Federal e municípios;
- ii) Contribuir para a redução da evasão, da reprovação e da distorção idade/série;
- iii) Oferecer atendimento educacional especializado às crianças, adolescentes e jovens com necessidades educacionais especiais;
- iv) Prevenir e combater o trabalho infantil, a exploração sexual e outras formas de violência contra crianças, adolescentes e jovens;
- v) Promover a formação da sensibilidade, da percepção e da expressão de crianças, adolescentes e jovens nas linguagens artísticas, literárias e estéticas;
- vi) Estimular crianças, adolescentes e jovens a manter uma interação efetiva em torno de práticas esportivas educacionais e de lazer;
- vii) Promover a aproximação entre a escola, as famílias e as comunidades; e,
- viii) Prestar assistência técnica e conceitual aos entes federados de modo a estimular novas tecnologias e capacidades para o desenvolvimento de projetos.

Para atender aos objetivos elencados, o programa estabelece uma articulação entre o Ministério da Educação (MEC), o Ministério do Desenvolvimento Social e Combate à Fome (MDS), o Ministério da Cultura (MinC) e o Ministério do Esporte (ME). Além disso, também conta com a participação de ações promovidas pelos estados, Distrito Federal, municípios e outras instituições públicas e privadas.

Nas escolas participantes do Mais Educação são oferecidas atividades optativas nos seguintes macrocampos: acompanhamento pedagógico, educação ambiental, esporte e lazer,

direitos humanos em educação, cultura e artes, cultura digital, promoção da saúde, comunicação e uso de mídias, investigação no campo das ciências da natureza e educação econômica.

Cada escola participante do programa deve desenvolver um projeto educativo no qual serão escolhidas seis atividades, a cada ano, com base nesses macrocampos, sendo que pelo menos uma destas atividades obrigatoriamente deve fazer parte do macrocampo de acompanhamento pedagógico. De acordo com o MEC (2013): "para o desenvolvimento de cada atividade, o governo federal repassa recursos para ressarcimento de monitores, aquisição dos kits de materiais, contratação de pequenos serviços e obtenção de materiais de consumo e permanentes."

A coordenação do programa é de responsabilidade da Secretaria de Educação Básica (SEB/MEC), em parceria com as Secretarias Estaduais e/ou Municipais de Educação. Sua operacionalização, todavia, é feita por meio do Programa Dinheiro Direto na Escola (PDDE) e pelo Programa Nacional de Alimentação Escolar (PNAE), que faz parte do Fundo Nacional de Desenvolvimento da Educação (FNDE). Por meio do PDDE, o governo presta assistência financeira às escolas públicas de educação básica e às escolas privadas de educação especial sem fins lucrativos, visando à melhoria de suas estruturas físicas e pedagógicas.

No que se refere às escolas atendidas, em 2008, foram cerca de 1.380, beneficiando 386 mil estudantes em 55 municípios espalhados pelo Brasil. Já em 2009, houve a ampliação para 5 mil escolas em 126 municípios, atendendo 1,5 milhão de estudantes. Nos anos de 2010 e 2011 houve um salto considerável no número de escolas atendidas: 10.027 e 14.995, respectivamente. Segundo o MEC (2013) a quantidade de estudantes beneficiados passou de 2,3 milhões, em 2010, para mais de 3 milhões em 2011.

Por fim, é importante ressaltar que os critérios de elegibilidade do programa não se mantiveram os mesmos de 2008 a 2011. Em 2008, o Programa adotava os seguintes critérios: escolas situadas em municípios que assinaram o compromisso Todos pela Educação; escolas com regularidade junto ao PDDE; escolas estaduais ou municipais localizadas nas capitais e cidades das regiões metropolitanas com mais de 200 mil habitantes, com baixo IDEB e com mais de 99 matrículas registradas no Censo 2007. Já em 2011 os critérios mudaram para: escolas contempladas com PDDE/Integral nos anos de 2008, 2009 e 2010; escolas de cidades com população igual ou superior a 18.844 habitantes que já fazem parte do PDE/Escola; escolas estaduais e/ou municipais que foram contempladas com o PDE/Escola 2007 e em 2009 ficaram com IDEB abaixo ou igual a 4,2 nas séries iniciais e/ou 3,8 nas séries finais.

Essas mudanças nos critérios de elegibilidade do Programa Mais Educação devem ser levadas em consideração para construção do grupo de controle (escolas que não aderiram ao



Programa) de modo que este represente um contrafactual adequado às escolas pertencentes ao grupo de tratamento.

### 3.3 FONTE E TRATAMENTO DOS DADOS

Para estimar os efeitos do Programa Mais Educação serão utilizados os microdados da Prova Brasil<sup>28</sup> de 2007, 2009 e 2011, o que permite a construção de um painel acompanhando as escolas públicas ao longo desses anos supracitados. Para identificar quais escolas foram atendidas pelo programa durante esse período, serão utilizadas as planilhas obtidas junto ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

A fim de tentar captar o efeito causal do programa, isto é, livre do efeito de outros fatores, são consideradas no modelo empírico as características das escolas, alunos, diretores, professores e características de localização geográfica, conforme descrito no Quadro 2. Não obstante, para mensurar o efeito do Mais Educação sobre a violência nas escolas, também serão utilizadas variáveis que representem adequadamente esse problema.

**Quadro 2 – Descrição das variáveis**

<b>Variáveis</b>	<b>Descrição das Variáveis de Impacto</b>
IDEB_5º_ano	Índice de Desenvolvimento da Educação Básica do 5º ano.
IDEB_9º_ano	Índice de Desenvolvimento da Educação Básica do 9º ano.
Nota_Média_Português_5º_ano	Nota médias das escolas do 5º ano em língua portuguesa.
Nota_Média_Matemática_5º_ano	Nota médias das escolas do 5º ano em matemática.
Nota_Média_Português_9º_ano	Nota médias das escolas do 9º ano em língua portuguesa.
Nota_Média_Matemática_9º_ano	Nota médias das escolas do 9º ano em matemática.
Taxa_de_abandono_5º_ano	Taxa de Abandono das Escolas do 5º ano.
Taxa_de_abandono_9º_ano	Taxa de Abandono das Escolas do 9º ano.
Índice_de_violência_5º_ano	Índice de violência das escolas para o 5º ano
Índice_de_violência_9º_ano	Índice de violência das escolas para o 9º ano
<b>Características dos Alunos e do Domicílio</b>	
Sexo_aluno_masculino	% de alunos do sexo masculino.
Raça_aluno_branco	% de alunos da raça branca.

<sup>28</sup> A Prova Brasil, oficialmente chamada de Avaliação Nacional do Rendimento Escolar (ANRESC), é uma avaliação censitária envolvendo alunos do 5º e 9º ano do Ensino Fundamental das escolas públicas das redes municipais, estaduais e federal do país.

Trabalho_aluno	% de alunos que trabalham fora de casa.
Incentivo_aluno	% de alunos cujos pais os incentivam a estudar.
Faz_tarefa_sempre	% de alunos que fazem a tarefa de casa sempre ou quase sempre.
Creche_ou_pré-escola	% de alunos que iniciaram os estudos na creche_ou_pré-escola.
Reprovação_aluno	% de alunos que nunca foram reprovados.
Mora_com_a_mãe	% de alunos que moram com a mãe.
Escol_mãe_ens_superior	% de alunos cuja a mãe completou a faculdade.
Prof_corrige_tarefa_sempre	% de alunos cujo professor corrige a tarefa de casa sempre ou quase sempre.
Televisão_aluno	% de alunos com televisão a cores em casa.
Geladeira_aluno	% de alunos com geladeira em casa.
Máquina_lavar_aluno	% de alunos com máquina de lavar em casa.
Computador_com_internet	% de alunos que têm computador com internet em casa.
Carro_aluno	% de alunos com carro em casa.
Banheiro_aluno	% de alunos com dois ou mais banheiros em casa.
Doméstica_aluno	% de alunos com empregada doméstica em casa.
<b>Características das Escolas</b>	
Telhado_escola_bom	1 se o estado de conservação do telhado da escola é considerado bom; 0 caso contrário.
Paredes_escola_bom	1 se o estado de conservação das paredes da escola é considerado bom; 0 caso contrário.
Pisos_escola_bom	1 se o estado de conservação dos pisos da escola é considerado bom; 0 caso contrário.
Banheiros_escola_bom	1 se o estado de conservação dos banheiros da escola é considerado bom; 0 caso contrário.
Hidráulica_escola_bom	1 se o estado de conservação das instalações hidráulicas da escola é considerado bom; 0 caso contrário.
Elétrica_escola_bom	1 se o estado de conservação das instalações elétricas da escola é considerado bom; 0 caso contrário.
Salas_escola_bom	1 se o estado de conservação das salas de aula da escola é considerado bom; 0 caso contrário.
Biblioteca_escola	1 se a escola possui biblioteca; 0 caso contrário.
Laboratório_escola	1 se a escola possui laboratório; 0 caso contrário.
Quadra_escola	1 se a escola possui quadra de esportes; 0 caso contrário.
Professores_vínculo_estável	1 se o percentual de professores com vínculo estável na escola for acima de 75%; 0 caso contrário.
<b>Características de localização</b>	

<i>Dummies_estaduais</i> <sup>29</sup>	Serão utilizadas <i>dummies</i> estaduais para captar a influência da localização geográfica sobre a probabilidade da escola ser atendida pelo Mais Educação.
População	População residente nos municípios.
<b>Características dos Professores</b>	
Sexo_professor_masculino	% de professores do sexo masculino.
Escol_prof_ens_superior	% de professores com ensino superior completo.
Experiência_professor	% de professores que leciona há 10 anos ou mais.
<b>Características do Diretor</b>	
Sexo_diretor_masculino	1 se o diretor da escola for do sexo masculino; 0 caso contrário.
Escol_diretor_ens_superior	1 se o diretor possui ensino superior em Pedagogia; 0 caso contrário.
Experiência_diretor	1 se o diretor trabalha em educação há mais de 15 anos; 0 caso contrário.
<b>Variáveis de Violência</b>	
Depredação_interna	1 se a escola apresenta sinais de depredação nas dependências internas; 0 caso contrário.
Depredação_externa	1 se a escola apresenta sinais de depredação nas dependências externas; 0 caso contrário.
Professor_atentado_vida	% de professores que alegaram terem sido vítimas de atentado à vida.
Professor_ameaçado_aluno	% de professores que alegaram terem sido ameaçados por algum aluno.
Professor_vítima_roubo	% de professores que alegaram terem sido vítimas de roubo por parte de algum aluno.
Alunos_bebidas	% de professores que alegaram que alunos frequentaram as aulas sob efeito de bebida alcoólica.
Alunos_drogas	% de professores que alegaram que alunos frequentaram as aulas sob efeito de drogas ilícitas.
Alunos_armas_branças	% de professores que alegaram que alunos frequentaram as aulas portando alguma arma branca.
Alunos_armas_fogo	% de professores que alegaram que alunos frequentaram as aulas portando alguma arma de fogo.
Gangues	1 se há a ocorrência de ação de gangues nas dependências externas da escola; 0 caso contrário.

Fonte: Microdados da Prova Brasil de 2007, 2009 e 2011.

Apresentada a fonte dos dados e a descrição das variáveis, a Tabela 7 contempla a média dos indicadores de impacto para os grupos de tratamento e controle. Como se observa, a taxa de crescimento do IDEB e das notas médias de proficiência em português e matemática das escolas contempladas pelo Mais Educação é superior as escolas do grupo de controle em quase

<sup>29</sup> Para as estimações do Nordeste a categoria base será o estado da Bahia, enquanto que para as estimações do Brasil, a categoria base será o Distrito Federal.

todos os níveis analisados, com exceções apenas para as escolas da Paraíba no 5º ano em 2009 e para os dados agregados do Brasil, no 5º ano em 2007.

Em relação às taxas de abandono, nota-se que estas diminuem com maior intensidade nos grupos de tratamento que contemplam as escolas de todo o país. Porém, ao considerar apenas as escolas do Nordeste, verifica-se uma redução maior nos grupos de controle, já com relação aos dados para o Estado da Paraíba, nota-se uma elevação na taxa de abandono no ano de 2009 tanto no grupo de tratamento quanto no grupo de controle. E, por fim, no caso do índice de violência percebe-se que entre as escolas do grupo de controle houve uma maior redução, com exceção às escolas do 9º ano de todo o Brasil.

Todavia, diante dessas estatísticas, não é possível afirmar que são decorrentes (ou não) do Programa Mais Educação, haja vista que podem existir outros fatores responsáveis por estas mudanças. É justamente essa a questão central que esse estudo busca responder.

Tabela 7 – Média dos indicadores de impacto para os grupos de controle e tratamento

IDEB													
	5º Ano						9º Ano						
	Paraíba		Nordeste		Brasil		Paraíba		Nordeste		Brasil		
	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	
2007	3,1447	3,0395	3,2229	3,1280	3,4077	3,4320	2,8005	2,7000	2,9216	2,8753	3,0774	3,1731	
2009	3,3121	3,3773	3,4746	3,3020	3,5950	3,3926	2,8808	2,7263	2,8655	2,9503	2,9762	3,1227	
2011	3,6248	3,5058	3,5844	3,4410	3,6643	3,4814	2,9936	2,7898	3,0173	2,9613	3,1495	3,1561	
Var. % (2007-2011)	15,27	15,34	11,22	10,01	7,53	1,44	6,89	3,33	3,28	2,99	2,34	-0,54	
Nota_Média_Português													
	5º Ano						9º Ano						
	Paraíba		Nordeste		Brasil		Paraíba		Nordeste		Brasil		
	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	
2007	160,6864	158,6747	157,4592	156,5538	160,8414	160,7156	220,7690	218,3357	215,7698	213,2726	220,7456	222,7116	
2009	167,4590	163,7517	165,0475	158,3740	167,2259	160,8871	231,4746	224,2205	227,3859	220,0235	229,7731	227,0202	
2011	167,1728	163,9415	164,8051	159,3893	167,3890	160,6965	225,1440	220,3182	223,1613	215,4199	227,5964	221,8868	
Var. % (2007-2011)	4,04	3,32	4,66	1,81	4,07	-0,01	1,98	0,91	3,43	1,01	3,10	-0,37	
Nota_Média_Matemática													
	5º Ano						9º Ano						
	Paraíba		Nordeste		Brasil		Paraíba		Nordeste		Brasil		
	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	Tratado	Controle	
2007	177,0342	175,5766	172,9868	172,3376	176,6735	177,0831	228,7939	226,9092	224,3767	222,6881	229,9669	233,0912	
2009	182,8877	180,7750	180,5170	174,2588	183,1131	177,1748	230,1832	226,2209	227,8301	221,9662	230,6337	229,3246	

<b>2011</b>	182,4212	179,2902	179,4961	174,7210	182,4646	176,0547	228,2752	227,2427	226,7101	220,6609	231,7861	227,5979
<b>Var. % (2007-2011)</b>	3,04	2,12	3,76	1,38	3,28	-0,58	-0,23	0,15	1,04	-0,91	0,79	-2,36
<b>Taxa_de_abandono</b>												
	<b>5º Ano</b>						<b>9º Ano</b>					
	<b>Paraíba</b>		<b>Nordeste</b>		<b>Brasil</b>		<b>Paraíba</b>		<b>Nordeste</b>		<b>Brasil</b>	
	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>
<b>2007</b>	8,4871	9,6840	5,8273	6,3868	4,4049	4,4450	13,9274	17,1746	12,1148	14,1102	9,7908	9,5051
<b>2009</b>	12,8178	13,0809	11,8496	10,1069	9,0591	8,8190	9,8551	12,8300	9,7090	8,6342	9,0757	7,5094
<b>2011</b>	6,2577	5,1304	4,3481	4,0772	3,7190	3,9467	9,3834	10,8176	7,5674	7,3966	6,6463	6,6452
<b>Var. % (2007-2011)</b>	-26,27	-47,02	-25,39	-36,16	-15,57	-11,21	-32,63	-37,01	-37,54	-47,58	-32,12	-30,09
<b>Índice_de_Violência</b>												
	<b>5º Ano</b>						<b>9º Ano</b>					
	<b>Paraíba</b>		<b>Nordeste</b>		<b>Brasil</b>		<b>Paraíba</b>		<b>Nordeste</b>		<b>Brasil</b>	
	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>	<b>Tratado</b>	<b>Controle</b>
<b>2007</b>	-1,0926	-1,0994	-0,6484	-0,7320	-0,7361	-0,8303	-0,6497	-0,8131	-0,7253	-0,9198	-0,8888	-1,1007
<b>2009</b>	-0,4531	-0,0448	0,5820	0,3473	0,4229	0,1921	-0,4065	0,3591	0,1770	-0,1992	0,0431	-0,4064
<b>2011</b>	-1,0969	-1,0847	-0,6847	-0,7814	-0,8316	-0,9504	-0,9117	-1,0922	-0,9642	-1,2993	-1,2121	-1,4462
<b>Var. % (2007-2011)</b>	-0,39	1,34	5,60	6,75	12,97	14,46	-40,33	-34,33	32,94	41,26	36,37	31,39

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2007, 2009 e 2011.

A Tabela 8, por sua vez, contém as estatísticas descritivas completas da amostra. Como pode ser observado, a maior parte dos alunos do 5º ano, tanto das escolas da Paraíba quanto do Nordeste e de todo o Brasil, são do sexo masculino. Já no 9º ano, ocorre o contrário.

Em todos os três casos, uma parcela considerável de estudantes afirmou possuir algum tipo de trabalho fora de casa e muitos alegaram que seus pais os incentivam a estudar. Um resultado que merece atenção é que são pouquíssimas as mães que possuem ensino superior completo.

Com relação às características das escolas, nota-se que mais da metade possui biblioteca e são poucas as que possuem laboratório. A maioria das escolas consideradas na amostra estão localizadas em Pernambuco e no Ceará, no modelo que considera apenas o Nordeste, e em São Paulo e Minas Gerais, no modelo que considera todo o país.

No que diz respeito às características dos professores e diretores, tem-se que ambos os cargos são desempenhados em sua maioria por indivíduos do sexo feminino com mais de 15 anos de experiência. Enquanto que a maior parte dos professores afirmaram possuir ensino superior completo, o mesmo não pode ser dito dos diretores.

Tabela 8 - Estatísticas descritivas dos alunos do 5º e 9º ano da Paraíba, do Nordeste e do Brasil

	Variáveis	Paraíba		Nordeste		Brasil	
		5º Ano	9º Ano	5º Ano	9º Ano	5º Ano	9º Ano
		Média	Média	Média	Média	Média	Média
<b>Variáveis de Impacto</b>	IDEB	3,3293	2,8180	3,3263	2,9464	3,4695	3,1250
	Nota_Média_Português	163,1179	222,6899	159,0670	218,2589	162,1782	224,4064
	Nota_Média_Matemática	179,3307	227,7040	174,5833	223,2024	178,0534	230,1054
	Taxa_de_abandono	9,0548	12,4402	6,6430	9,4447	5,2943	7,9740
	Índice_de_Violência	-0,9309	-0,6247	-0,4996	-0,8906	-0,6561	-1,0724
<b>Características do Aluno e do Domicílio</b>	Sexo_aluno_masculino	0,5036	0,43404	0,5149	0,4336	0,5121	0,4509
	Raça_aluno_branco	0,2941	0,2691	0,2475	0,2046	0,2706	0,2684
	Trabalho_aluno	0,1893	0,1801	0,2114	0,2183	0,1914	0,2124
	Incentivo_aluno	0,9440	0,9795	0,9400	0,9759	0,9485	0,9799
	Faz_tarefa_sempre	0,7220	0,6324	0,7125	0,6389	0,7065	0,5938
	Creche_ou_pré-escola	0,7207	0,8244	0,7001	0,7791	0,6916	0,7649
	Reprovação_aluno	0,5298	0,5000	0,5394	0,5274	0,5626	0,5675
	Mora_com_a_mãe	0,8883	0,8786	0,8756	0,8670	0,8783	0,8737
	Escol_mãe_ens_superior	0,0642	0,0416	0,0787	0,0458	0,0868	0,0516
	Prof_corrige_tarefa_sempre	0,7839	0,8278	0,7732	0,8402	0,7639	0,8223
	Situação_econômica_aluno	-0,0005	-0,0512	0,0545	0,0400	0,0794	0,0451
<b>Características da Escola</b>	Estrutura_escola	-0,1966	-0,1772	-0,2755	-0,2779	-0,2774	-0,2724
	Biblioteca_escola	0,6277	0,8128	0,6066	0,7666	0,6450	0,8108
	Quadra_escola	0,3105	0,5312	0,2750	0,4780	0,4065	0,6550
	Laboratório_escola	0,1519	0,4033	0,1618	0,3313	0,1941	0,4010
	Professores_vínculo_estável	0,5603	0,4383	0,5970	0,5742	0,5762	0,5308
<b>Características de localização</b>	População	227.750	255.758	350.878	371.079	561.403	862.403
	d_Alagoas	-	-	0,0515	0,0471	0,0290	0,0211
	d_Ceará	-	-	0,1444	0,1744	0,0813	0,0781
	d_Maranhão	-	-	0,1362	0,1337	0,0766	0,0599
	d_Paraíba	-	-	0,0642	0,0592	0,0361	0,0265
	d_Pernambuco	-	-	0,1741	0,1773	0,0980	0,0794



	d_Piauí	-	-	0,0402	0,0337	0,0226	0,0151
	d_Rio_Grande_Norte	-	-	0,0610	0,0593	0,0343	0,0265
	d_Sergipe	-	-	0,0421	0,0394	0,0237	0,0176
	d_Rondônia	-	-	-	-	0,0108	0,0134
	d_Acre	-	-	-	-	0,0031	0,0020
	d_Amazonas	-	-	-	-	0,0343	0,0257
	d_Roraima	-	-	-	-	0,0022	0,0029
	d_Pará	-	-	-	-	0,0779	0,0404
	d_Amapá	-	-	-	-	0,0075	0,0046
	d_Tocantins	-	-	-	-	0,0053	0,0042
	d_Minas_Gerais	-	-	-	-	0,0333	0,0684
	d_Espírito_Santo	-	-	-	-	0,0150	0,0202
	d_Rio_de_Janeiro	-	-	-	-	0,0848	0,0961
	d_São_Paulo	-	-	-	-	0,0408	0,0884
	d_Paraná	-	-	-	-	0,0131	0,0281
	d_Santa_Catarina	-	-	-	-	0,0078	0,0097
	d_Rio_Grande_do_Sul	-	-	-	-	0,0479	0,0740
	d_Mato_Grosso_do_Sul	-	-	-	-	0,0152	0,0194
	d_Mato_Grosso	-	-	-	-	0,0115	0,0097
	d_Goiás	-	-	-	-	0,0249	0,0384
	d_Bahia	-	-	-	-	0,1612	0,1235
<b>Características do Professor</b>	Sexo_professor_masculino	0,1216	0,3999	0,1113	0,3505	0,1202	0,3087
	Escol_prof_ens_superior	0,3571	0,9745	0,7295	0,9228	0,7523	0,9528
	Experiência_professor	0,8234	0,6290	0,6751	0,6375	0,6700	0,6363
<b>Características do Diretor</b>	Sexo_diretor_masculino	0,0762	0,2268	0,1567	0,2552	0,1663	0,2454
	Escol_diretor_ens_superior	0,8084	0,2755	0,3819	0,2925	0,4068	0,3168
	Experiência_diretor	0,7927	0,8798	0,7378	0,7646	0,7673	0,8134
<b>Número de Observações</b>		1.053	657	16.404	11.099	29.146	24.787

Fonte: Elaborada pelo autor a partir dos Microdados da Prova Brasil de 2007, 2009 e 2011.

### **3.4 ESTRATÉGIA EMPÍRICA**

Nesta seção será apresentada a estratégia empírica adotada neste estudo. Inicialmente, serão apresentados os critérios de elegibilidade do Mais Educação adotados na construção dos grupos de controle e tratamento. A avaliação de impacto será realizada através da combinação de dois métodos: o método das Diferenças em Diferenças e o método de Pareamento por Escore de Propensão.

O procedimento de Análise de Componentes Principais será empregado para a construção do índice de violência utilizado como indicador de impacto, bem como para a construção de dois índices a serem usados como variáveis explicativas: um índice de estrutura física utilizando as variáveis referentes às condições dos telhados, paredes, pisos, banheiros e instalações hidráulicas e elétricas das escolas, e um índice de condição econômica do domicílio dos estudantes através das variáveis que indicam a posse de televisão, geladeira, máquina de lavar, computador com acesso à internet, carro, número de banheiros e presença de empregada doméstica no domicílio.

#### **3.4.1 GRUPOS DE TRATAMENTO E DE CONTROLE**

Para construção dos grupos de tratamento e controle, foram empregados os critérios de elegibilidade das escolas pelo Mais Educação para o ano de 2011. Conforme tais critérios, e visando a construção de um grupo de controle consistente e equilibrado, excluiu-se da amostra as escolas federais, as escolas situadas em municípios com população inferior a 18.844 habitantes e as escolas com IDEB superior a 4,2 (5º ano) e 3,8 (9º ano)<sup>30</sup>.

#### **3.4.2 MODELO EMPÍRICO: DIFERENÇAS EM DIFERENÇAS**

O modelo de Diferenças em Diferenças (DD) é um procedimento bastante empregado na literatura de avaliação de políticas públicas e busca isolar o efeito de um determinado programa ou projeto social sob uma dada variável de impacto (Y). De modo geral, o que o estimador de DD faz é calcular uma dupla diferença de médias: a diferença entre as médias da variável de impacto para os grupos de tratamento e controle antes e depois do programa, e a diferença entre essas diferenças.

---

<sup>30</sup> Tais critérios podem ser consultados no Manual da Educação Integral para o exercício de 2011, disponível no portal do MEC.

O modelo pode ser aplicado tanto para dados em painel quanto para dados de seção cruzada repetidos no tempo. O importante é que haja informação disponível para as unidades tratadas e não tratadas em ao menos um período anterior e posterior ao programa. A principal vantagem do DD é permitir o controle pelas características não observáveis, assumidas como sendo invariantes no tempo, e que são uma das fontes mais preocupantes do viés de seleção. De modo mais formal, usando a notação de Foguel (2012), o estimador de DD pode ser calculado por meio das seguintes formas equivalentes:

$$DD = \{E[Y|T = 1, t = 1] - E[Y|T = 1, t = 0]\} - \{E[Y|T = 0, t = 1] - E[Y|T = 0, t = 0]\} \quad (9)$$

$$DD = \{E[Y|T = 1, t = 1] - E[Y|T = 0, t = 1]\} - \{E[Y|T = 1, t = 0] - E[Y|T = 0, t = 0]\} \quad (10)$$

Nas equações acima,  $T = 1$  indica a presença do tratamento, enquanto que  $T = 0$  indica a ausência, e  $t = 1$  denota o período posterior ao início do programa, enquanto que  $t = 0$  denota o período anterior. Como ressaltado por Gertler et al. (2011), a primeira diferença da equação (9) controla os fatores constantes no tempo, uma vez que o mesmo grupo de tratamento está sendo avaliado nos dois períodos. A segunda diferença, por sua vez, permite capturar os fatores externos que variam no tempo, já que o grupo de controle esteve sujeito ao mesmo ambiente econômico que o grupo de tratamento. Dessa forma, a diferença entre a primeira e a segunda diferença fornecerá o impacto da intervenção livre de viés, isto é, sem superestimar ou subestimar o verdadeiro efeito do programa.

Outra forma de obter o estimador de DD é através de regressões lineares. Seguindo Foguel (2012), considere o seguinte modelo para a unidade  $i$  no tempo  $t$ :

$$Y_{it} = \delta X_{it} + \gamma T_{it} + \rho t_{it} + \beta(T_{it}t_{it}) + \varepsilon_{it} \quad (11)$$

Em que:  $X$  contempla as características observáveis das unidades consideradas,  $T$  é uma variável binária que assume valor 1 se a unidade recebe o tratamento e 0 caso contrário,  $t$  é outra variável binária que assume 1 no período posterior ao programa e 0 caso contrário, e  $\varepsilon$  é o termo de erro estocástico. O efeito médio do tratamento é captado pelo parâmetro  $\beta$  associado às variáveis interagidas, sob a hipótese de que  $E[\varepsilon|X, T, t] = 0$ .

Conforme apontado por Khandker, Koolwal e Samad (2010), também é possível obter o estimador de DD empregando o modelo de efeitos fixos para dados em painel. Esse

procedimento é adequado, principalmente, em análises que envolvem múltiplos períodos de tempo. Para tanto, considere o modelo abaixo:

$$Y_{it} = \delta X_{it} + \phi T_{it} + \eta_i + \varepsilon_{it} \quad (12)$$

Onde  $\eta$  representa a heterogeneidade individual, considerada invariante no tempo, que pode ser correlacionada com qualquer característica observada e com  $\varepsilon_{it}$ . Diferenciando ambos os lados da equação (12), obtém-se:

$$(Y_{it} - Y_{it-1}) = \delta(X_{it} - X_{it-1}) + \phi(T_{it} - T_{it-1}) + (\eta_i - \eta_i) + (\varepsilon_{it} - \varepsilon_{it-1}) \quad (13)$$

A equação (13) pode ser reescrita como:

$$\Delta Y_{it} = \delta \Delta X_{it} + \phi \Delta T_{it} + \Delta \varepsilon_{it} \quad (14)$$

Como pode ser observado, a heterogeneidade individual ( $\eta$ ) é cancelada na diferenciação. Assim, é possível estimar a equação (13) pelo método dos Mínimos Quadrados Ordinários (MQO), já que o problema de endogeneidade não está mais presente. A magnitude não viesada do efeito do tratamento é dado pelo parâmetro  $\phi$ , que será idêntico à estimativa de DD da equação (10) para o caso em que considera-se apenas dois períodos de tempo (KHANDKER, KOOLWAL E SAMAD, 2010).

Apesar de sua vasta aplicação e simplicidade, DD requer hipóteses mais fortes e restritivas que os métodos de aleatorização<sup>31</sup>. Inicialmente, como de praxe nas avaliações de impacto, espera-se que o grupo de controle represente um contrafactual adequado para o grupo de tratamento. Em outras palavras, espera-se que as unidades não tratadas sejam o mais similar possível das unidades tratadas, de modo que elas possam representar o que teria ocorrido com estas últimas na ausência do tratamento.

Outra condição do modelo de DD, conforme apontado por Foguel (2012), é que não ocorra qualquer mudança após o início da intervenção capaz de afetar de forma heterogênea os grupos de tratamento e controle. Se esta última condição não for obedecida, o estimador irá erroneamente atribuir os efeitos dessa mudança ao impacto do programa.

---

<sup>31</sup> Métodos baseados em critérios de seleção aleatória para a formação dos grupos de tratamento e controle.

Por fim, a hipótese mais crítica do DD é a chamada hipótese de mesma tendência temporal (*Equal Trend Assumption*), também conhecida como hipótese de tendências paralelas (*Parallel-Trend Assumption*). Segundo Gertler et al. (2011), a referida hipótese afirma que na ausência do tratamento, a variável de impacto teria de crescer ou decrescer à mesma taxa, tanto para o grupo de tratamento quanto para o grupo de controle. Assim, a principal limitação do método das Diferenças em Diferenças é a sua sensibilidade em relação aos fatores que podem influenciar os grupos de forma heterogênea, o que pode invalidar ou viesar a estimativa do impacto (GERTLER et al., 2011).

As hipóteses citadas acima têm por objetivo mimetizar os métodos de aleatorização, de modo a tornar a análise o mais próximo possível de um método experimental. Para construir um grupo de controle mais consistente e representativo do grupo de tratamento é possível combinar DD com o método de Pareamento por Escore de Propensão<sup>32</sup> (*Propensity Score Matching* – PSM). De acordo com Khandker, Koolwal e Samad (2010), o procedimento consiste em realizar o PSM no período pré-tratamento para parear as unidades tratadas com as unidades não tratadas que forem similares em um conjunto de atributos observáveis selecionados pelo pesquisador. Com isso, será criada uma região de suporte comum contendo as unidades que serão usadas para estimar o DD. O estimador obtido dessa maneira é mais robusto e eficiente que o estimador puro de DD, conforme verificado por Hirano, Imbens e Ridder (2003). Sendo assim, este será o procedimento adotado nesta pesquisa.

Como critério de pareamento das escolas será empregado o método de *Kernel*, onde as unidades tratadas são pareadas com a média ponderada das unidades de controles, sendo os pesos dessa ponderação inversamente proporcionais à diferença entre o escore de propensão das unidades tratadas e não tratadas.

Ao considerar três períodos de tempo, como é o caso desse estudo, é preciso adaptar o modelo para o que vem a ser conhecido como DD em dois estágios. Desse modo, a regressão a ser estimada pode ser descrita por:

$$Y_i = \alpha + \beta_1 T_{2009i} + \beta_2 T_{2011i} + \gamma_1 t_{2009i} + \gamma_2 t_{2011i} + \delta_1 (T_{2009i} t_{2009i}) + \delta_2 (T_{2011i} t_{2011i}) + \theta X_{it} + \varepsilon_i \quad (15)$$

Em que:  $i = 1 \dots N$  representa as N escolas;  $T_{2009i} = 1$  indica a presença do programa em 2009, enquanto que  $T_{2009i} = 0$  indica a ausência;  $T_{2011i} = 1$  indica a presença do programa em 2011, enquanto que  $T_{2011i} = 0$  indica a ausência;  $t_{2009i} = 1$  se o ano for 2009, enquanto que  $t_{2009i} =$

---

<sup>32</sup> Tal combinação foi proposta por Heckman, Ichimura e Todd (1998).

0 em caso contrário;  $t_{2011i} = 1$  se o ano for 2011, enquanto que  $t_{2011i} = 0$  em caso contrário;  $X_{it}$  é uma matriz com as características dos alunos, do domicílio, das escolas, dos professores e dos diretores; e  $\varepsilon_i$  é o termo de erro estocástico. O efeito total do programa é captado pela soma dos parâmetros estimados  $\delta_1$  e  $\delta_2$ .

### 3.4.3 ANÁLISE DE COMPONENTES PRINCIPAIS

O método de Análise de Componentes Principais (ACP) tem por objetivo explicar a estrutura de variância-covariância de um conjunto de variáveis aleatórias através de combinações lineares dessas mesmas variáveis. Tais combinações lineares são chamadas de componentes principais e são não correlacionadas entre si (JOHNSON e WICHERN, 2007).

Para entender como a ACP funciona, considere um sistema com  $p$  variáveis aleatórias. Para reproduzir a variabilidade total desse sistema seria necessário a construção de  $p$  componentes principais. Todavia, grande parte dessa variabilidade pode ser adequadamente contabilizada por um número  $k < p$  de componentes. Dessa forma, os  $k$  componentes obtidos podem substituir a informação contida nas  $p$  variáveis originais, simplificando a análise.

Seguindo a notação de Johnson e Wichern (2007), a ACP pode ser formalmente abordada considerando um vetor aleatório  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$ , com matriz de variância-covariância  $\Sigma_{p \times p}$ . Sejam  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$  os autovalores de  $\Sigma_{p \times p}$  e  $e_1, e_2 \dots e_p$  os respectivos autovetores normalizados, os componentes principais são as combinações lineares dadas por:

$$\begin{aligned} Y_1 &= \mathbf{e}_1' \mathbf{X} = e_{11}X_1 + e_{12}X_2 + \dots + e_{1p}X_p \\ Y_2 &= \mathbf{e}_2' \mathbf{X} = e_{21}X_1 + e_{22}X_2 + \dots + e_{2p}X_p \\ &\vdots \\ Y_p &= \mathbf{e}_p' \mathbf{X} = e_{p1}X_1 + e_{p2}X_2 + \dots + e_{pp}X_p \end{aligned} \tag{16}$$

Em que:  $e_i' e_i = 1 \quad \forall i = 1, 2, \dots, p$ ;

$e_i' e_k = 0 \quad \forall i \neq k$  (condição de ortogonalidade dos vetores);

$var(Y_i) = \mathbf{e}_i' \Sigma \mathbf{e}_i = \lambda_i \quad \forall i = 1, 2, \dots, p$ ;

$cov(Y_i, Y_k) = \mathbf{e}_i' \Sigma \mathbf{e}_k = 0 \quad \forall i \neq k$ .

As combinações lineares são definidas de modo que o primeiro componente tenha a maior variância, o segundo componente a segunda maior variância e assim por diante.

É possível calcular a proporção da variância total de  $\mathbf{X}$  explicada pelo  $i$ -ésimo componente principal através da seguinte fórmula:

$$\frac{var(Y_i)}{var(\mathbf{X})} = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} = \frac{\lambda_i}{tr(\Sigma_{p \times p})} \quad (17)$$

De modo similar, a proporção da variância total explicada por  $k$  componentes principais é dada por:

$$\frac{\sum_{i=1}^k var(Y_i)}{var(\mathbf{X})} = \frac{\sum_{i=1}^k \lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} = \frac{\sum_{i=1}^k \lambda_i}{tr(\Sigma_{p \times p})} \quad (18)$$

Segundo Johnson e Wichern (2007), se cerca de 80 a 90% da variância total puder ser atribuída a um número  $k$  de componentes, então esses componentes podem substituir as  $p$  variáveis originais sem perda de informação relevante.

Vale ressaltar, todavia, que a ACP pura não é muito adequada quando se utiliza variáveis dicotômicas, como é o caso dessa pesquisa. Desse modo, será utilizada uma matriz de correlações do tipo *polychoric* para a obtenção dos componentes, seguindo o procedimento descrito por Kolenikov e Angeles (2004). Tal procedimento assume que cada variável binária é uma manifestação de variáveis contínuas subjacentes e estima a ACP com base nessa nova matriz de correlações.

Como dito anteriormente, a ACP será empregada para a construção de três índices: um índice de violência que será utilizado como variável de impacto para captar os efeitos do programa Mais Educação na redução da violência nas escolas, um índice de estrutura física e um índice de condição econômica do domicílio dos discentes, ambos usados como variáveis explicativas. Seguindo Filmer e Pritchett (2001), Baker and Wurgler (2006), Abramovay et. al. (2013), entre outros, os índices corresponderão aos primeiros componentes obtidos em cada ACP realizada.

### 3.5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Esta seção contempla a análise e discussão dos resultados e tem como objetivo mensurar o efeito do programa Mais Educação sobre as escolas do 5º e 9º ano na Paraíba, na região Nordeste e no Brasil, considerando como variáveis de impacto o IDEB, a nota média das escolas em português e em matemática, a taxa de abandono e o índice de violência. Vale

ressaltar que a intenção em obter estimativas para a Paraíba, para o Nordeste e para todo o país é verificar se os resultados mantêm o mesmo padrão em todos os casos. Ou seja, busca-se averiguar se o Programa apresenta efeitos similares quando consideradas as diferentes delimitações geográficas.

A seguir são apresentados os resultados da combinação de DD com PSM de *Kernel*. Os escores de propensão estimados, bem como os gráficos de densidade que mostram se a hipótese de sobreposição foi mantida, podem ser consultados no apêndice.

### 3.5.1 IMPACTOS SOBRE O IDEB

A partir dos resultados expostos na Tabela 9, observa-se que apenas as escolas do 5º ano, tanto para o Nordeste quanto para todo o Brasil, apresentam resultados estatisticamente significativos. Portanto, pode-se dizer que, de maneira agregada, o Programa Mais Educação reduz o abandono e/ou repetência, aumenta a taxa de aprovação, e/ou aumenta a proficiência dos alunos. Essa análise decorre do fato do IDEB resultar da combinação entre as pontuações em testes padronizados e informações sobre rendimento escolar (taxa de aprovação, repetência e evasão).

De maneira geral, pode-se dizer que o efeito estimado corrobora com os alguns dos objetivos do programa, isto é, suas atividades tendem a impactar diretamente em fatores chaves do IDEB. Todavia, vale ressaltar que se faz apropriado analisar os impactos sobre os indicadores de forma desagregada.

**Tabela 9 – Estimação dos efeitos do Mais Educação sobre o IDEB**

	Paraíba		Nordeste		Brasil	
	5º Ano	9º Ano	5º Ano	9º Ano	5º Ano	9º Ano
<b>IDEB</b>	0,166 (0,124)	0,111 (0,158)	0,038* (0,022)	0,008 (0,035)	0,139*** (0,019)	0,002 (0,021)
<b>Escolas no Grupo de Tratamento</b>	386	219	3.911	2.465	8.766	6.347
<b>Escolas no Grupo de Controle</b>	189	100	4.110	2.485	7.363	6.323

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007, 2009 e 2011.

Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%. Erros padrões robustos entre parênteses.



### 3.5.2 IMPACTOS SOBRE AS NOTAS MÉDIAS DAS ESCOLAS

Pela Tabela 10, nota-se que o Programa Mais Educação apresenta efeito positivo sobre a proficiência média em português das escolas do 5º ano no Brasil. Por outro lado, no caso do 9º ano, observa-se um resultado contrário ao esperado no que diz respeito aos impactos sobre as escolas do Nordeste e do Brasil, uma vez que aquelas atendidas pelo programa apresentam efeitos negativos sobre a proficiência em língua portuguesa.

Em relação a proficiência em matemática, os resultados foram significativos apenas para o 9º ano. Sendo assim, pode-se dizer que a escola que aderiu o Mais Educação apresenta resultado superior em relação às escolas que não aderiram.

De modo geral, não se pode concluir ao certo se o referido programa apresenta bons resultados no que se refere às notas médias de proficiências das escolas, haja vista que pelos resultados estimados, o Mais Educação afeta sobretudo o desempenho no 9º ano, porém de maneira negativa em português e positiva em matemática. Cabe salientar aqui que a verificação do impacto em relação à proficiência foi apenas um exercício, dado que melhoria nos indicadores de nota não estão explicitamente elencados entre os objetivos do programa<sup>33</sup>.

Além disso, é importante ressaltar que esses resultados não corroboram com os encontrados por Xerxenevsky (2012), uma vez que a autora encontrou efeito nulo para os alunos do 9º ano do Rio Grande do Sul e efeito positivo em português para os alunos do 5º ano.

**Tabela 10 – Estimação dos efeitos do Mais Educação sobre a média de português e matemática das escolas**

	Paraíba		Nordeste		Brasil	
	5º Ano	9º Ano	5º Ano	9º Ano	5º Ano	9º Ano
<b>Português</b>	-1,425 (3,693)	-3,302 (4,456)	0,679 (0,532)	-3,537*** (0,884)	1,770*** (0,446)	-4,142*** (0,540)
<b>Matemática</b>	-1,427 (3,094)	4,123 (4,742)	-0,451 (0,542)	1,948** (0,805)	0,773 (0,492)	1,417*** (0,534)
<b>Escolas no Grupo de Tratamento</b>	386	216	3.911	2.455	8.766	6.326
<b>Escolas no Grupo de Controle</b>	189	97	4.110	2.473	7.363	6.317

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007, 2009 e 2011.

Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%. Erros padrões robustos entre parênteses.

<sup>33</sup> Vide página 45 desta dissertação.

### 3.5.3 IMPACTOS SOBRE A TAXA DE ABANDONO

A redução da taxa de abandono consiste em um objetivo explícito do Mais Educação, talvez por esse ser um dos principais problemas da educação brasileira, além de ser bastante oneroso. Diante dos resultados, pode-se dizer que o programa apresenta efeito de redução desse problema, com exceção para as escolas do 9º ano da região Nordeste (efeito positivo). Sendo assim, de uma maneira geral, pode-se dizer que o Programa alcança um de seus principais objetivos, que consiste em reduzir as taxas de abandono.

A título de comparação, verifica-se que esses resultados são semelhantes aos apresentados por Pereira (2011), e a intuição por trás dos mesmos parece ser a seguinte. Para os alunos do 5º ano, que são mais jovens, atividades extracurriculares que proporcionem maior acesso aos meios de cultura, bem como atividades esportivas e artísticas, tornam a escola mais atraente e menos cansativa, o que pode estimular esses alunos a não abandoná-las vis-à-vis os alunos do 9º ano, geralmente já na adolescência, que podem atribuir uma maior importância ao tempo disponível extraclasse para a realização de outras atividades

**Tabela 11 – Estimação dos efeitos do Mais Educação sobre a taxa de abandono**

	Paraíba		Nordeste		Brasil	
	5º Ano	9º Ano	5º Ano	9º Ano	5º Ano	9º Ano
<b>Taxa de Abandono</b>	-10,827*** (2,251)	-1,579 (2,984)	-5,263*** (0,351)	1,379** (0,613)	-5,430*** (0,287)	0,178 (0,325)
<b>Escolas no Grupo de Tratamento</b>	379	214	3.859	2.430	8.679	6.290
<b>Escolas no Grupo de Controle</b>	185	98	4.068	2.454	7.301	6.269

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007, 2009 e 2011.

Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%. Erros padrões robustos entre parênteses.

### 3.5.4 IMPACTOS SOBRE O ÍNDICE DE VIOLÊNCIA

No que diz respeito aos efeitos do Mais Educação sobre a violência nas escolas, os resultados apresentam maior parcimônia, além de convergir para um dos objetivos do programa, que consiste na sua redução.

Observa-se na Tabela 12, que para as escolas do 5º e do 9º ano, tanto no Nordeste quanto no Brasil como um todo, além de na Paraíba para as escolas do 5º ano, o programa apresentou impactos negativos, isto é, de redução da violência.

Os efeitos encontrados nessa etapa podem ser explicados pelo próprio desenho do programa, que ao manter crianças e jovens mais tempo nas escolas desempenhando atividades esportivas, de acompanhamento pedagógico, etc., fazem com que as mesmas estejam menos vulneráveis às situações de violência. Além disso, também acredita-se que essas mesmas atividades conscientizem os alunos a manter distância de drogas ilícitas e lícitas, conscientizando para a importância dos estudos.

**Tabela 12 – Estimação dos efeitos do Mais Educação sobre o índice de violência**

	Paraíba		Nordeste		Brasil	
	5º Ano	9º Ano	5º Ano	9º Ano	5º Ano	9º Ano
<b>Índice de Violência</b>	-4,709*** (1,160)	-0,546 (0,534)	-1,549*** (0,395)	-0,958*** (0,356)	-1,454*** (0,373)	-0,930*** (0,183)
<b>Escolas no Grupo de Tratamento</b>	84	42	925	630	2.360	1.830
<b>Escolas no Grupo de Controle</b>	21	8	809	494	1.749	1.552

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007, 2009 e 2011.

Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%. Erros padrões robustos entre parênteses.

### 3.6 CONSIDERAÇÕES FINAIS

Essa pesquisa se propôs a avaliar os impactos do Programa Federal Mais Educação nas escolas do Estado da Paraíba, da região Nordeste e do Brasil, sobre os seguintes indicadores: IDEB, proficiência média em português e matemática, taxa de abandono e um índice de violência.

Para isto, utilizou-se os microdados da Prova Brasil de 2007, 2009 e 2011, juntamente com os arquivos obtidos junto ao INEP para identificar quais escolas, em cada ano, foram atendidas pelo programa. Dessa maneira, foi possível acompanhar as escolas em um período anterior (2007), e em dois períodos posteriores (2009 e 2011) ao programa.

Pela estrutura dos dados, optou-se por montar um painel das escolas. Sendo assim, a estratégia empírica para estimar os impactos do Mais Educação consistiu na combinação do modelo de Diferenças em Diferenças com o PSM a partir do método de *Kernel*.

A partir da análise dos resultados pode-se dizer que o Programa Mais Educação apresenta efeitos positivos sobre o IDEB do 5º ano das escolas do Nordeste e do Brasil. O programa também apresenta efeito sobre a proficiência média de matemática das escolas do 9º ano. Contudo, o mesmo não pode ser dito acerca dos seus efeitos sobre a proficiência média de português.

Em relação à redução das taxas de abandono, foram encontrados efeitos significativos e de considerável magnitude para o 5º ano. Entretanto, no caso do 9º ano, não se observa nenhum efeito de redução sobre as taxas de abandonos das escolas na Paraíba, no Nordeste e no Brasil.

Por fim, no que se refere à questão da violência nas escolas, os resultados indicam que o programa contribui significativamente para a redução desse problema em todos os casos considerados.

É importante ressaltar que muitos dos objetivos do programa Mais Educação não são captados explicitamente pelos indicadores supracitados, uma vez que esse programa tem como objetivo central fornecer uma formação mais completa e cidadã aos estudantes através da expansão da jornada escolar. Para tanto, uma série de atividades, nos mais variados macrocampos, são realizadas, o que torna o programa complexo e com alguns resultados difíceis de serem mensurados. Sendo assim, os resultados dessa pesquisa não permite concluir definitivamente se o programa deve ou não continuar sendo empregado. Para chegar a essa resposta seria preciso, além de utilizar uma série de outras variáveis de impacto, calcular o retorno econômico do programa e verificar se seus benéficos superam os custos<sup>34</sup>.

Todavia, o presente estudo oferece uma reflexão para o papel que um programa governamental tenha no que se refere à queda das taxas de abandono, dado que o referido problema é um dos maiores gargalos da educação brasileira.

---

<sup>34</sup> Esta etapa não foi realizada nesse estudo porque não foi possível obter as informações necessárias para tanto. Todavia, fica como proposta para futuras pesquisas.

## REFERÊNCIAS

- ABRAMOVAY, R.; RODRIGUES JÚNIOR, M.; MADEIRA, G. A.; GONÇALVES, M. F.; MACIEL, I. S. R.; SANTOS, R. A. **Cinco anos de Agroamigo: retrato do público e efeitos do programa**. 2. ed. Fortaleza: Banco do Nordeste, 2013.
- BAKER, M.; WUGLER, J. Investor Sentiment and the Cross-Section of Stock Returns. **The Journal of Finance**, v. 61, n. 4, p. 1645-1680, 2006.
- BARROS, R. P.; LIMA, L. Avaliação de Impacto de Programas Sociais: por que, para que e quando fazer? In: MENEZES FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012. p. 13-29.
- BRASIL. Decreto nº 7.083, de 27 de janeiro de 2010. Dispõe sobre o Programa Mais Educação. **Diário Oficial da União**, Brasília, 27 de janeiro de 2010.
- BRASIL. Portaria Normativa Interministerial nº 17, de 24 de abril de 2007. Institui o Programa Mais Educação, que visa fomentar a educação integral de crianças, adolescentes e jovens, por meio do apoio a atividades sócio-educativas no contraturno escolar. **Diário Oficial da União**, Brasília, 26 de abril de 2007.
- FILMER, D.; PRITCHETT, L. H. Estimating wealth effects without expenditure data-or-tears: na application to educational enrollments in States of India. **Demography**, v. 38, n. 1, p. 115-132, 2001.
- FOGUEL, M. N. Diferenças em Diferenças. In: MENEZES FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012. p. 69-83.
- GARCIA, S. M. C. **Política de Educação Integral: avaliação do programa Mais Educação no sistema público municipal de ensino de Fortaleza**. 144 f. Dissertação (Mestrado) – Universidade Federal do Ceará, Pró-Reitoria de Pesquisa e Pós-Graduação, Mestrado Profissional em Avaliação de Políticas Públicas, Fortaleza, 2013.
- GERTLER, P. J.; MARTINEZ, S.; PREMAND, P.; RAWLINGS, L. B.; VERMEERSCH, C. M. J. **Impact Evaluation in Practice**. Washington, DC: The World Bank, 2011.
- GRUBER, J. **Finanças Públicas e Política Pública**. Rio de Janeiro: LTC, 2009.
- HECKMAN, J.; ICHIMURA, H.; TODD, P. Matching as an econometric evaluation estimator. **Review of Economic Studies**, v. 65, p. 261-294, 1998.
- HIRANO, K.; IMBENS, G. W.; RIDDER, G. Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. **Econometrica**, v. 71, n. 4, p. 1161–1189, 2003.
- JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. Upper Saddle River, NJ: Pearson Prentice Hall, 2007.
- KHANDKER, S. R.; KOOLWAL, G. B.; SAMAD, H. A. **Handbook on impact evaluation: Quantitative Methods and Practices**. Washington, DC: The World Bank, 2010.

KOLENIKOV, S.; ANGELES, G. The Use of Discrete Data in PCA: Theory, Simulations, and Applications to Socioeconomic Indices. **CPC/MEASURE Working Paper**, n. WP-04-85.

LOPES, L. C.; ARAÚJO, J. C. Avaliação do Programa Mais Educação na Rede Municipal de Fortaleza: a realidade das escolas da regional V. In: XVI ENDIPE - Encontro Nacional de Didática e Práticas de Ensino, 2012, Campinas. E-books do XVI ENDIPE – Encontro Nacional de Didática e Práticas de Ensino - ARARAQUARA - SP: Junqueira&Marin Editores, 2012.

MENDES, K. D. **O Impacto do Programa Mais Educação no Desempenho dos Alunos da Rede Pública Brasileira**. 2011. 57 f. Trabalho de Conclusão de Curso (Graduação). Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2011.

NÓBREGA, S. A.; SILVA, M. A. Programa Mais Educação: ampliando quantidades, garantindo qualidade? In: XXVI Simpósio Brasileiro de Políticas e Administração da Educação, 2013, Recife. Política, Planos e Gestão da Educação: democratização e qualidade social, 2013. v. 17.

PEREIRA, G. C. **Uma Avaliação de Impacto do Programa Mais Educação no Ensino Fundamental**. 2011. 96 f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Instituto de Economia, 2011.

PIRES, A. M. S.; AGUIAR, D. S.; SANTANA, F. D.; OLIVEIRA, H. A.; SILVA, K. M. P.; ROSA, A. C. F. Análise do Programa Mais Educação em Paraíso do Tocantins: um estudo de caso no colégio estadual Idalina de Paula. In: VII CONNEPI - Ações Sustentáveis para o desenvolvimento Regional, 2012, Palmas - TO.

MENEZES FILHO, N. (Org.). **Avaliação econômica de projetos sociais**. 1. ed. São Paulo: Dinâmica Gráfica e Editora, 2012.

MINISTÉRIO DA EDUCAÇÃO. **Programa Mais Educação passo a passo**. Brasília: MEC, SECAD, 2009. Disponível em:

< [http://portal.mec.gov.br/dmdocuments/passoapasso\\_maiseducacao.pdf](http://portal.mec.gov.br/dmdocuments/passoapasso_maiseducacao.pdf) > Acesso em: 07/07/2014.

\_\_\_\_\_. **Saiba Mais - Programa Mais Educação**. MEC, 2013. Disponível em: < [http://portal.mec.gov.br/index.php?option=com\\_content&id=16689&Itemid=1115](http://portal.mec.gov.br/index.php?option=com_content&id=16689&Itemid=1115) > Acesso em 09/07/2014.

XERXENEVSKY, L. L. **Programa Mais Educação: avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul**. 2012. 142 f. Dissertação (Mestrado em Economia do Desenvolvimento) – Faculdade de Administração, Contabilidade e Economia, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.

## APÊNDICE

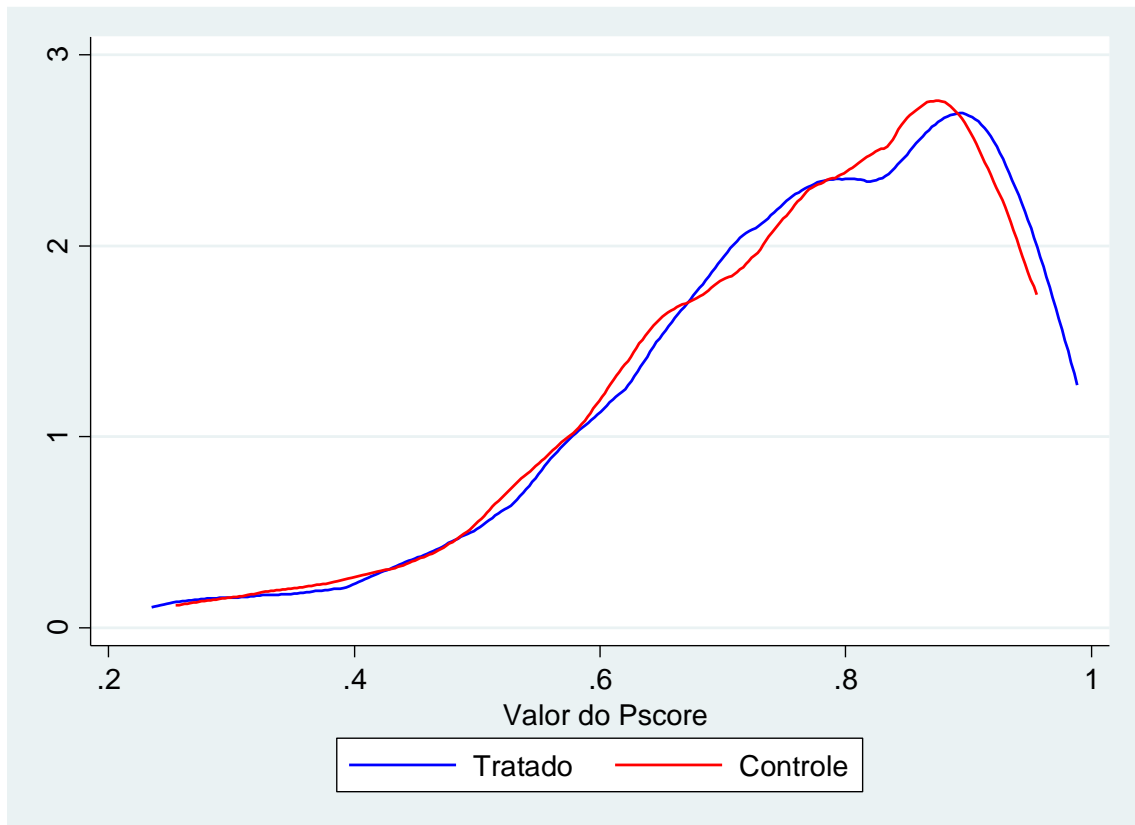
Tabela 13 – Estimação do escore de propensão para as escolas da Paraíba

	5º Ano		9º Ano	
	Coefficiente	Erro-Padrão	Coefficiente	Erro-Padrão
Estrutura	0,0003	0,0784	-0,0998	0,1304
Biblioteca_escola	0,1267	0,2561	-0,9824*	0,5151
Quadra_escola	0,0503	0,2718	-0,3268	0,3744
Laboratório_escola	0,1199	0,4185	0,0360	0,3921
Sexo_aluno_masculino	2,0068*	1,1528	-0,4822	1,5658
Raça_aluno_branco	0,0666	0,9606	1,9152	1,8384
Condição_econômica_aluno	0,1912**	0,1019	0,0725	0,1591
Mora_com_a_mãe	1,3542	1,5664	-2,2254	2,3351
Escol_mãe_ens_superior	-1,5499	2,6676	2,3302	5,3602
Incentivo_aluno	2,7041	2,3450	-9,4258	6,3303
Trabalho_aluno	1,6372	1,4206	-1,5887	1,8515
Creche_ou_pré-escola	-0,5595	0,9599	0,0634	2,0202
Reprovação_aluno	0,7960	0,8922	2,5669*	1,5271
Faz_tarefa_sempre	0,9486	1,1246	-2,2624	1,7538
Prof_corrige_tarefa_sempre	-0,7832	0,9738	1,7341	1,4017
Sexo_diretor_masculino	0,3213	0,3866	0,0798	0,4035
Escol_diretor_ens_superior	0,0983	0,2497	0,6095	0,4204
Experiência_diretor	0,5469*	0,3192	-0,5297	0,7363
Professores_vínculo_estável	-0,5147**	0,2486	-0,1677	0,3684
Sexo_professor_masculino	-0,9990*	0,5558	0,3856	0,6538
Escol_professor_ens_superior	-0,6823*	0,3605	0,4828	1,5817
Experiência_professor	-0,4543	0,3745	0,3723	0,5727
População	1,92E-06***	5,86E-07	2,15E-06*	8,77E-07
Intercepto	-3,7074	2,6859	10,7827	6,9930

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007 e 2011.

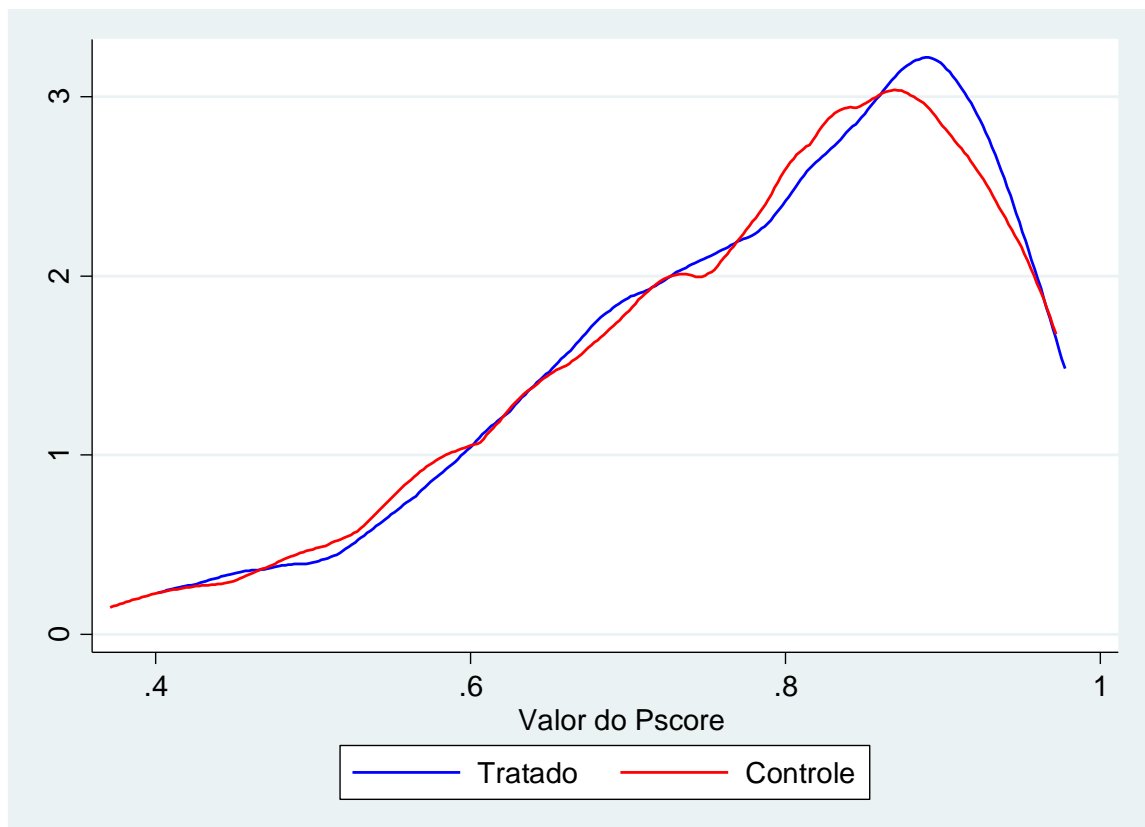
Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%.

**Figura 1 – Função de Densidade do Escore de Propensão (Paraíba - 5º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011.

**Figura 2 – Função de Densidade do Escore de Propensão (Paraíba - 9º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011



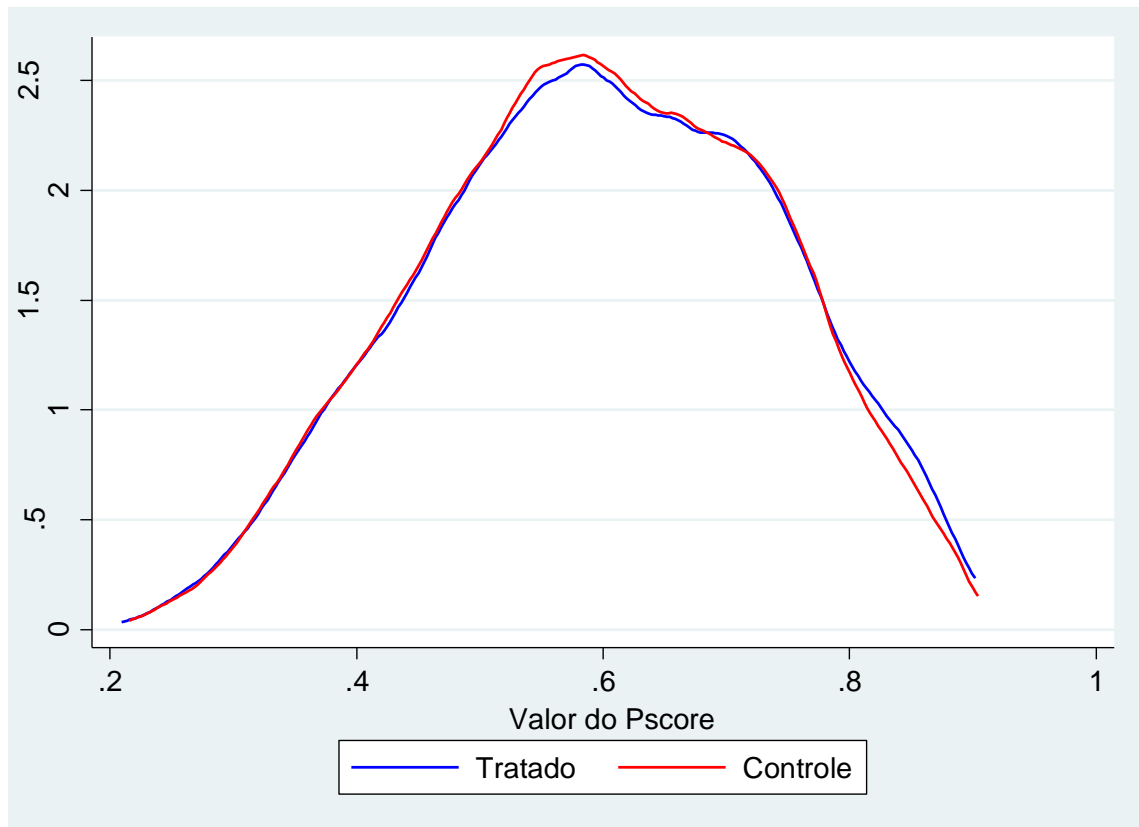
**Tabela 14 – Estimação do escore de propensão para as escolas do Nordeste**

	5º Ano		9º Ano	
	Coefficiente	Erro-Padrão	Coefficiente	Erro-Padrão
Estrutura	0,0144	0,0182	0,0201	0,0269
Biblioteca_escola	0,0398	0,0605	-0,2813***	0,0960
Quadra_escola	0,1081*	0,0649	0,0352	0,0827
Laboratório_escola	0,1217	0,0947	0,0040	0,0933
Sexo_aluno_masculino	0,1142	0,2900	1,0538**	0,4090
Raça_aluno_branco	-0,7424***	0,2481	-0,4005	0,4218
Condição_econômica_aluno	0,1578***	0,0229	0,1733***	0,0365
Mora_com_a_mãe	0,4242	0,3816	-0,6274	0,5340
Escol_mãe_ens_superior	-0,0981	0,5395	-4,1372***	1,0420
Incentivo_aluno	0,2141	0,4988	-1,3591	1,3203
Trabalho_aluno	-0,3209	0,3131	-0,9839**	0,3835
Creche_ou`pré-escola	-0,1654	0,2093	0,4392	0,3693
Reprovação_aluno	0,5798***	0,2097	0,9771***	0,3178
Faz_tarefa_sempre	-0,2559	0,2594	0,3147	0,3819
Prof_corrige_tarefa_sempre	-0,8850***	0,2403	-0,2907	0,3736
Sexo_diretor_masculino	0,0345	0,0800	-0,1770*	0,0939
Escol_diretor_ens_superior	0,0605	0,0573	0,0404	0,0882
Experiência_diretor	0,1947***	0,0694	0,0634	0,1120
Professores_vínculo_estável	-0,0389	0,0579	0,0601	0,0825
Sexo_professor_masculino	0,0325	0,1196	-0,2185	0,1334
Escol_professor_ens_superior	-0,1095	0,0727	0,2913	0,2084
Experiência_professor	0,0533	0,0717	-0,2197*	0,1214
População	3,93E-07***	4,65E-08	8,65E-07***	8,60E-08
d_Alagoas	-0,2147*	0,1262	-0,9909***	0,1876
d_Ceará	0,7947***	0,1060	0,3670**	0,1570
d_Maranhão	0,5086***	0,1006	0,1466	0,1426
d_Paraíba	1,2017***	0,1288	0,8422***	0,1983
d_Pernambuco	0,5607***	0,0914	-0,0714	0,1384
d_Piauí	0,5650***	0,1355	0,3237	0,2286
d_Rio_Grande_do_Norte	0,7204***	0,1365	0,2049	0,1908
d_Sergipe	-0,0025	0,1445	-0,6833***	0,2045
Intercepto	-0,1893	0,5903	1,1579	1,3447

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007 e 2011.

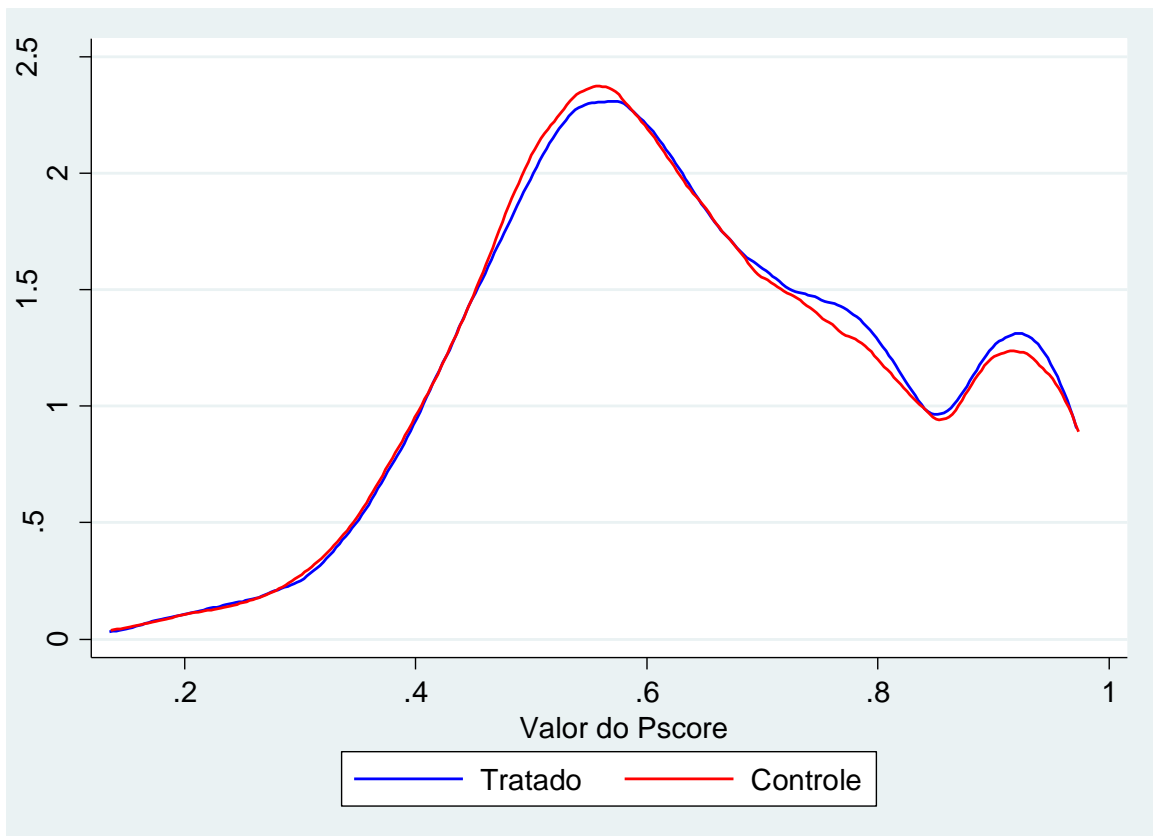
Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%.

**Figura 3 – Função de Densidade do Escore de Propensão (Nordeste - 5º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011

**Figura 4 – Função de Densidade do Escore de Propensão (Nordeste - 9º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011

**Tabela 15 – Estimação do escore de propensão para as escolas do Brasil**

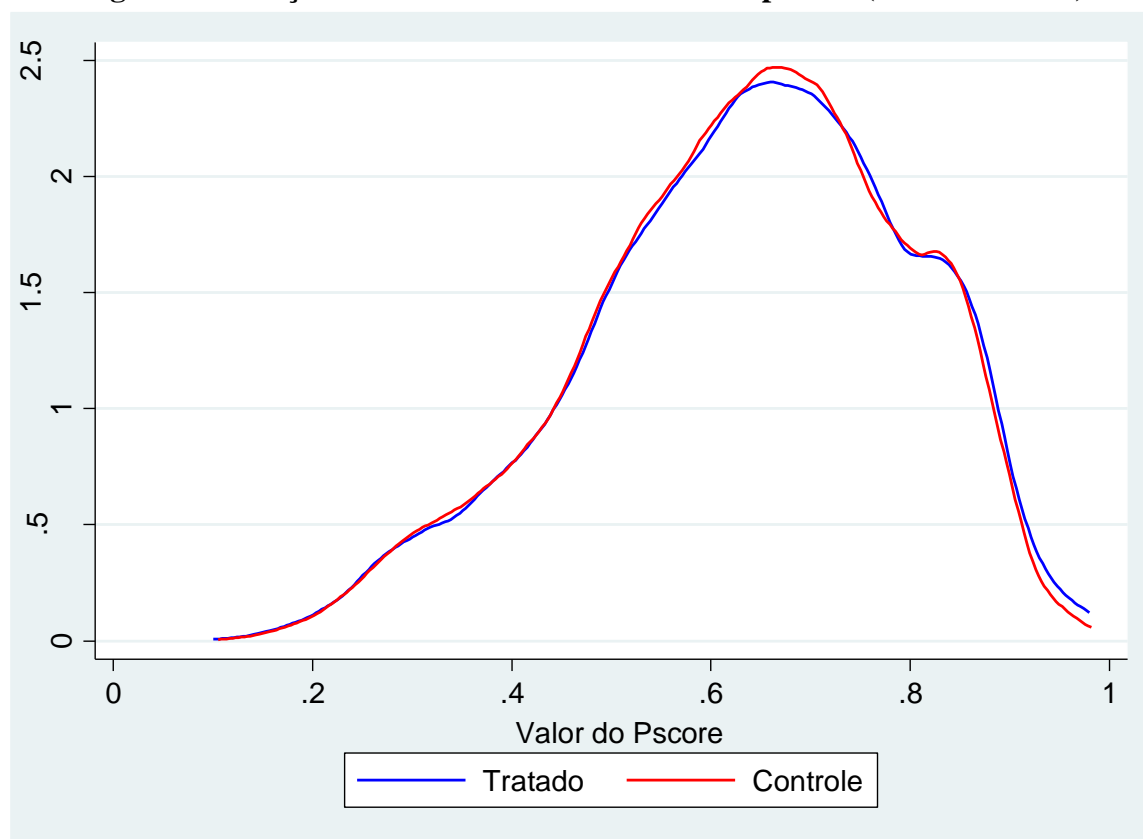
	5º Ano		9º Ano	
	Coefficiente	Erro-Padrão	Coefficiente	Erro-Padrão
Estrutura	-0,0001	0,0127	-0,0070	0,0164
Biblioteca_escola	0,0219	0,0448	-0,2035***	0,0716
Quadra_escola	0,1080**	0,0449	0,1048*	0,0593
Laboratório_escola	0,3171***	0,0579	0,2224***	0,0550
Sexo_aluno	0,0576	0,2007	1,0429***	0,2725
Raça_aluno_branco	-1,1126***	0,1737	-1,9708***	0,2467
Condição_econômica_aluno	0,2340***	0,0185	0,1833***	0,0270
Mora_com_a_mãe	0,0515	0,2629	-1,4507***	0,3599
Escol_mãe_ens_superior	-0,0420	0,3470	-6,0023***	0,6237
Incentivo_aluno	0,0846	0,4022	-1,5651	1,0095
Trabalho_aluno	-0,5741**	0,2335	-1,5528***	0,2520
Creche_ou_pré-escola	-0,9977***	0,1475	-0,1823	0,2252
Reprovação_aluno	0,3124**	0,1526	-0,0483	0,2077
Faz_tarefa_sempre	-0,2937	0,1804	-0,1162	0,2336
Prof_corrige_tarefa_sempre	-0,8062***	0,1698	-0,5528**	0,2305
Sexo_diretor	-0,0049	0,0547	-0,1565***	0,0589
Escol_diretor_ens_superior	0,0929**	0,0398	0,0090	0,0547
Experiência_diretor	0,1158**	0,0525	0,0897	0,0776
Professores_vínculo_estável	-0,0251	0,0405	0,0245	0,0526
Sexo_professor	-0,1442***	0,0831	0,1183	0,0874
Escol_prof_ens_superior	-0,0203	0,0563	0,6585***	0,1778
Experiência_professor	0,0130	0,0522	-0,1363*	0,0766
População	2,38E-08*	1,27E-08	-2,06E-08	1,33E-08
d_Alagoas	-1,2207***	0,4396	-2,1111***	0,4700
d_Ceará	-0,0529	0,4328	-0,4820	0,4623
d_Maranhão	-0,5134	0,4333	-1,0982**	0,4623
d_Paraíba	0,2517	0,4395	-0,1869	0,4756
d_Pernambuco	-0,3593	0,4308	-1,0842**	0,4559
d_Piauí	-0,4553	0,4413	-0,7554	0,4917
d_Rio_Grande_do_Norte	-0,2493	0,4404	-0,7253	0,4704
d_Sergipe	-1,0091**	0,4443	-1,7990***	0,4808
d_Rondônia	-0,2788	0,4614	-0,3570	0,5028
d_Acre	0,2922	0,5487	0,1925	0,8895
d_Amazonas	-0,5604	0,4397	-0,6012	0,4743
d_Roraima	-	-	1,3600	1,1228
d_Pará	-0,0963	0,4313	-0,0444	0,4661
d_Amapá	1,5632**	0,6265	1,0187	0,7476
d_Tocantins	-0,0816	0,4815	0,2645	0,6366
d_Minas_Gerais	-1,1140**	0,4318	-1,6504***	0,4519
d_Espírito_Santo	-0,9213**	0,4442	-1,0240**	0,4736
d_Rio_de_Janeiro	0,3681	0,4313	-0,3791	0,4508
d_São_Paulo	-2,4277***	0,4319	-2,4036***	0,4555

d_Paraná	-1,2927***	0,4440	-1,1056**	0,4653
d_Santa_Catarina	-1,2536***	0,4495	-1,4835***	0,4833
d_Rio_Grande_do_Sul	-0,5393	0,4342	-0,5367	0,4595
d_Mato_Grosso_do_Sul	-2,1673***	0,4491	-2,8615***	0,4899
d_Mato_Grosso	-0,9121**	0,4499	-0,8973*	0,5173
d_Goiás	-0,8472*	0,4338	-0,3529	0,4632
d_Bahia	-0,9361**	0,4287	-0,9365*	0,4535
Intercepto	2,3497***	0,6294	4,9899***	1,1029

Fonte: Estimativas obtidas a partir dos microdados da Prova Brasil de 2007 e 2011.

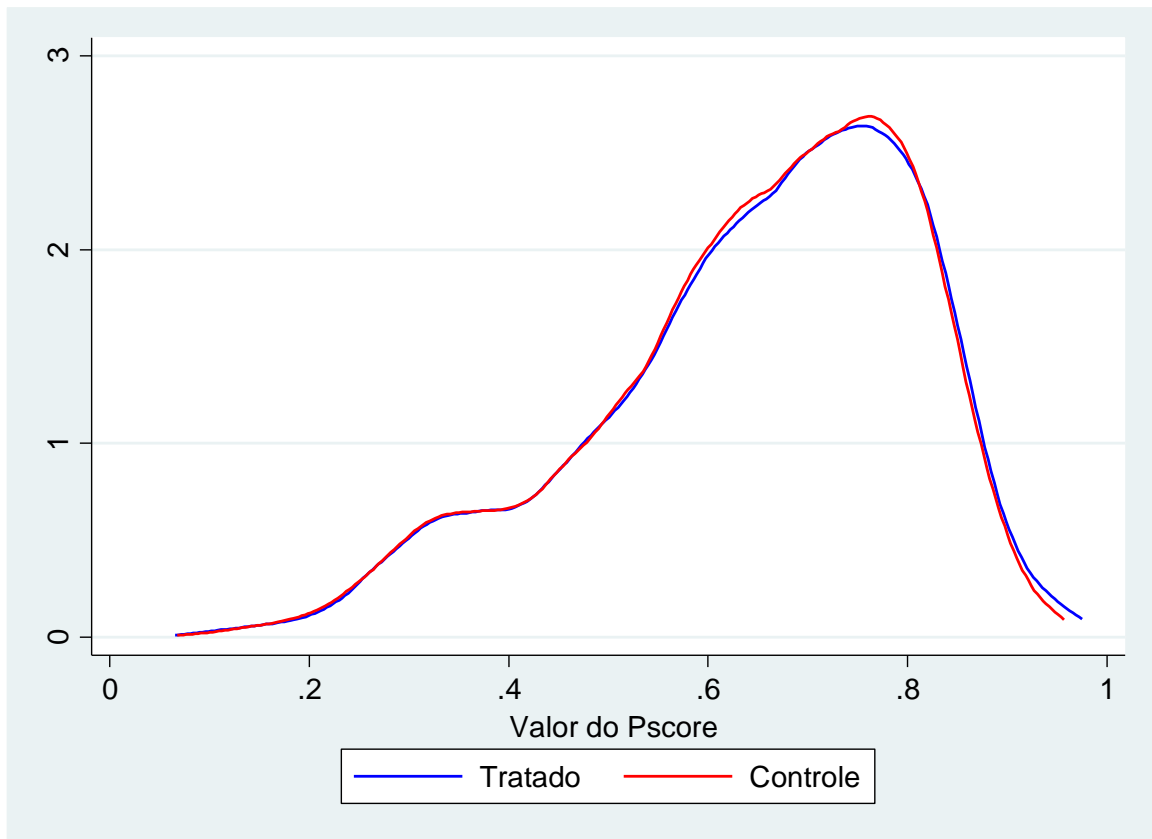
Nota: \*Estatisticamente significativa a 10%; \*\*Estatisticamente significativa a 5%; \*\*\*Estatisticamente significativa a 1%. A *dummy* referente ao estado de Roraima foi excluída da estimação do 5º ano por causa de problema de multicolinearidade.

**Figura 5 – Função de Densidade do Escore de Propensão (Brasil - 5º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011

**Figura 6 – Função de Densidade do Escore de Propensão (Brasil - 9º Ano)**



Fonte: Elaboração própria a partir dos microdados da Prova Brasil de 2007 e 2011